UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE FÍSICA

Identification and quantification of the alveolar compartment by Confocal Laser Endomicroscopy in patients with Interstitial Lung Diseases

Rita Gonçalves Frazão da Rocha Pinto

Mestrado Integrado em Engenharia Biomédica e Biofísica
Perfil de Biofísica Médica e Fisiologia de Sistemas

Dissertação orientada por:
Prof. Dr. med. Jouke T. Annema
Prof. Dr. Nuno Matela

2018

AMDG.
Para a minha família.

ACKNOWLEDGEMENTS

The work presented in these pages involved the contribution and support of several people to which I am very grateful.

Thank you, Lizzy and Paul, for accepting me as your student and for the countless hours you have spent meeting with me. Thank you for your outstanding dedication and kindness. Thanks Lizzy for your patience while teaching me so much about the research world and showing me all the clinical procedures. Thanks Paul for being a true engineer and for all your out-of-the-box ideas.

Thank you Dr. Jouke Annema for your passion for innovative imaging techniques, your dedication to this project, the trust you have placed in my work and the contagious enthusiasm about every new result.

Thank you Prof. Nuno Matela for the willingness to discuss my work, for the guidance, motivation and such relevant feedback which was crucial for these pages to become a thesis.

Thanks to the research group from the Department of Respiratory Medicine at the AMC who welcomed me for 6 months. Thank you for all the knowledge you shared with me at the journal club and group meetings. Thank you for joining me in my lunch breaks and for the kindness in every conversation. Particularly, thanks to those who shared the room F5-2(.0)60 with me, generously giving me a desk inside one of the most crowded rooms in the hospital.

Thank you to the 20 patients which altruistically accepted to be subject to the CLE procedure and to give their data to this project. I hope this work contributes to the Interstitial Lung Diseases research. Also, thanks to all the medical and hospital staff involved in the data acquisition. Thank you to the multidisciplinary team who discussed these patients' cases. Particularly, thanks to the pathologist Dr. Joris Roelofs for discussing with us the findings in the histology and CLE images.

Thank you to the Mauna Kea Technologies (MKT) for creating such an interesting technique. Also, for welcoming us and gathering a team to discuss our project.

Thank you, Toby Cornish and Knut Kvaal, for replying to my emails and sharing your work with me. Your ImageJ plugin "GLCM_Texture_Too v0.009" was very useful for me to develop the texture analysis algorithm used in this project.

Thanks to the good friends who stood by my side and gave me emotional, spiritual and scientific support. Thank you to my friends in Amsterdam, particularly, Helena and Niloo. Thank you to each one of my Fixolas friends who shared this adventure with me and were always present even being hundreds of kilometers away. Thank you Gambozinos, my friends in the Lord.

Thank you, Vasco, for motivating me and for your unconditional love and friendship.

Finally, thanks to my family for being the biggest supporters (and funders) of this work. Thank you Mãe, Pai, avô Zé, avó Violete, Carolina, Maria, Francisco, Luísa, António and the little ones Francisca, Francisco Xavier, Joana e Tomás for all the love and patience. Thanks to Tia Lisa and Tio Vasco, who have always believed (and overestimated) the potential of my skills and talents.

ABSTRACT

Interstitial Lung Diseases (ILD) is a heterogeneous group of more than 200 diseases which affect the lung parenchyma. To identify the type of ILD a patient suffers from is a difficult process, and 10% of the patients are categorized as unclassifiable, mostly due to the absence of histopathological data associated with the risks of lung biopsies. The patient specific diagnosis is important because of its implications to the patient treatment and management, being particularly relevant to identify lung fibrosis.

The Confocal Laser Endomicroscopy (CLE) can add information to this process. CLE allows to image the lung tissue with a micrometer resolution in a minimally invasive way, through a bronchoscopy. The elastin fibers from the lung alveoli are visible with this technique due to their autofluorescence. Since there is evidence that the amount of elastin fibers increases, and their architecture is altered in lung fibrosis, CLE should be used to extract values reflecting this condition.

Thus, the main goal of this project was to improve the CLE technique and increase its usability, by extracting numerical values from the images which would reflect the state of the alveolar space, particularly the elastin fibers.

The ILD patients recruited for the study had their lung alveoli imaged with CLE. The CLE movies were selected, pre-processed – were converted into frames, had their image quality enhanced and some mosaics were obtained – and then analyzed.

The ridge detection algorithm detected most fibers recognized by a human observer. It allowed the measurement of the Number of Detected Fibers, their Length and Width, the Number of Junctions between fibers and to calculate the Sum from all Fibers' Lengths. The Gray-Level Co-occurrence Matrix allowed the extraction of the Haralick texture features: Angular Second Moment (Energy), Entropy, Inverse Difference Moment, Contrast, Variance and Correlation. These algorithms produced consistent and unbiased numerical features, in an efficient process which can analyze the entire data set in a few seconds.

Regarding the fiber related measurements, it was expected for the fibrotic patients to have wider fibers and a higher number of fibers and junctions. In terms of texture variables, it was expected from the fibrotic patients to present higher values of Entropy, Contrast and Variance, and lower values of Inverse Difference Moment, given their lung tissue should correspond to more complex and heterogeneous images with more ridges present.

Due to the small sample size, it was still not possible to stratify patients with this data set. Nevertheless, the measurements presented here already contribute to the study of ILD, helping to understand the disease physiology. It is hoped that in the future, these measurements will aid the diagnosis process specially in those cases when patients cannot undergo a surgical biopsy. Additionally, CLE could potentially be used as an anti-fibrotic medication efficiency measurement tool.

Keywords: Confocal Laser Endomicroscopy, Interstitial Lung Diseases, Morphometric Analysis, Texture Analysis, Elastin Fibers.

RESUMO

Doenças Intersticiais Pulmonares (DIP) é um termo que inclui mais de 200 doenças que afectam o parênquima pulmonar, partilhando manifestações clínicas, radiográficas e patológicas semelhantes. Este conjunto de doenças é bastante heterogéneo, apresentando cada tipo de DIP em diferente grau os elementos de inflamação e fibrose: enquanto a inflamação é reflectida pelo aumento de células inflamatórias e presença de nódulos ou edema, a fibrose reflecte-se pelas fibras adicionais de colagénio e elastina.

Identificar o tipo de DIP de um doente é um processo difícil, sendo a Discussão Multidisciplinar o actual método de diagnóstico "*gold standard*": vários médicos especialistas compõem uma equipa multidisciplinar que vai ter em conta os dados clínicos, radiológicos e patológicos disponíveis para chegar a uma conclusão. Estes dados incluem imagens de tomografia computorizada de alta resolução (TCAR), a descrição da lavagem broncoalveolar e, quando possível, dados de biópsias. Apesar do esforço e competência da equipa multidisciplinar, 10% dos pacientes são categorizados como inclassificáveis devido a dados inadequados ou discrepância entre os dados existentes. A maior causa para DIP inclassificáveis é a ausência de dados histopatológicos associada aos riscos das biópsias cirúrgicas.

É muito importante determinar a DIP específica de um doente, dadas as suas implicações no tratamento e gestão do mesmo. É particularmente crítica a distinção entre doentes com Fibrose Pulmonar Idiopática (FPI) e doentes sem FPI, dado que há terapias anti-fibróticas – como o *Pirfenidone* – indicadas para FPI que são extremamente dispendiosas, exigindo certeza no diagnóstico antes de serem prescritas. Além disso, o tratamento com agentes imunossupressores pode funcionar com o grupo dos não-FPI mas aumenta a morte e hospitalizações nos doentes com FPI.

A discussão multidisciplinar pode beneficiar da informação adicional oferecida pelo *Confocal Laser Endomicroscopy* (CLE), uma técnica de imagiologia que torna possível visualizar os alvéolos pulmonares com resolução microscópica de forma minimamente invasiva, através de uma broncoscopia. O laser do CLE tem um comprimento de onda de 488 nm que permite observar a autofluorescência das fibras de elastina. Há evidências de que a quantidade de fibras de elastina é aumentada e a arquitectura destas fibras é alterada na presença de fibrose pulmonar, a qual está associada a algumas doenças intersticiais pulmonares incluindo a fibrose pulmonar idiopática.

Até à data, os vídeos de *Confocal Laser Endomicroscopy* são, na maioria dos casos, analisados apenas visualmente, e pouca informação objectiva e consistente foi conseguida destes vídeos em doentes de DIP. No entanto, é possível obter informação mais relevante dos mesmos, convertendo-os em *frames*, pré-processando as imagens e extraindo atributos numéricos.

Neste projecto, foram obtidas imagens dos alvéolos pulmonares de doentes de DIP através de CLE. O principal objectivo do projecto é melhorar a técnica de CLE e aumentar a sua usabilidade para que no futuro possa contribuir para facilitar a estratificação de doentes com DIP e eventualmente reduzir o número de biópsias pulmonares nestes doentes.

Como mencionado, o instrumento de *Confocal Laser Endomicroscopy* emite uma luz laser azul de 488nm, a qual é reflectida no tecido e reorientada para o sistema de detecção pela mesma lente, passando por um pequeno orifício (*pinhole*). Isto permite que a luz focada seja recolhida e que feixes provenientes de planos fora de foco sejam excluídos, originando uma resolução microscópica que permite imagens ao nível celular. Quando o CLE é aplicado a imagem pulmonar, é possível observar as paredes alveolares pela autofluorescência natural presente nas fibras de elastina.

No estudo clínico subjacente a este estudo, o protocolo de CLE foi aplicado a 20 pacientes, embora 8 tenham sido posteriormente excluídos da análise.

Os vídeos de CLE obtidos sofreram duas selecções: uma com base na região onde uma biópsia (usada como referência) foi tirada e outra com base na qualidade técnica das imagens. Depois, os dados foram pré-processados: geraram-se imagens mosaico com um campo de visão alargado e, paralelamente converteram-se as sequências de vídeo em *frames*. A qualidade da imagem foi melhorada, filtrando o ruído electrónico para que posteriormente pudesse ser aplicada a análise de imagem. Esta análise extraiu valores numéricos que reflectem o estado do espaço alveolar, nomeadamente, variáveis de textura e medições relacionadas com as fibras de elastina.

As imagens de CLE obtidas mostraram-se muito interessantes. A resolução é superior à tomografia computorizada de alta resolução e a tridimensionalidade acrescenta informação às biópsias. O facto de permitir feedback em tempo real e observar ao vivo os movimentos naturais da respiração contribui para a análise do estado do doente.

A análise de textura feita às imagens serviu-se de um algoritmo de extracção de variáveis de Haralick a partir de uma *Gray-Level Co-occurence Matrix* (GLCM). Foram extraídas as variáveis de textura Momento Angular Secundário (Energia), Entropia, Momento de Diferença Inversa, Contraste, Variação e Correlação.

O algoritmo de *Ridge Detection* (detecção de linhas) identificou a maior parte das fibras de elastina detectáveis por um observador humano e mediu o Número de Fibras, o seu Comprimento e Largura e o Número de Junções entre fibras, permitindo também calcular a Soma dos Comprimentos de todas as fibras.

Estes algoritmos devolveram valores consistentes num processo mais eficiente comparado com um observador humano, conseguindo avaliar em poucos segundos múltiplas variáveis para todo o conjunto de dados.

As medições relacionadas com as fibras de elastina pretendiam ajudar a identificar os doentes fibróticos. Era esperado que as fibras dos doentes fibróticos fossem mais largas, mas isso não se observou. Também se previa que este grupo de doentes apresentasse maior número de fibras e junções, mas não houve uma diferença significativa entre grupos. No entanto, quando o grupo fibrótico foi segregado, o número de fibras e junções parece separar a fibrose moderada da fibrose severa. Este resultado é interessante na medida em que sugere que a monitorização do número de fibras/junções com CLE pode potencialmente ser usado como medida de eficácia de medicação anti-fibrótica.

Em relação às variáveis de textura, esperava-se que os doentes fibróticos apresentassem valores mais elevados de Entropia, Contraste e Variância e valores inferiores de Momento de Diferença Inversa, dado que o seu tecido pulmonar deveria corresponder a imagens mais complexas e heterogéneas com mais arestas presentes. No entanto, ainda não foi possível estabelecer diferenças significativas entre grupos.

Apesar dos resultados com o conjunto de dados usado não ter demonstrado correlações fortes entre as conclusões do CLE e da TCAR/histopatologia, os valores das variáveis em si já contribuem para o estudo das DIP, nomeadamente da sua fisiologia. De facto, a amostra de doentes deste estudo era reduzida, mas com uma amostra maior, espera-se que algumas das varáveis se correlacionem com outras técnicas usadas no diagnóstico e permitam segregar os pacientes em grupos e eventualmente aplicar classificação de dados. Neste momento, é possível especular que algumas variáveis seriam melhores candidatas para um classificador, nomeadamente os Números de Fibras e Junções, a Soma dos Comprimentos das fibras e as variáveis de Haralick Entropia e Energia.

O projecto apresentado nesta dissertação foi desenvolvido através de um estágio de 6 meses no departamento de Pneumologia no Academic Medical Center em Amsterdão, Países Baixos. No Academic Medical Center (AMC), fui acompanhada pelos estudantes de doutoramento Lizzy Wijmans - médica - e Paul Brinkman - engenheiro biomédico - e supervisionada pelo Dr. Jouke Annema, MD, PhD, Professor de endoscopia pulmonar. Este grupo de investigação do AMC está focado em técnicas inovadoras de imagiologia do sistema pulmonar e teve a oportunidade de reunir com a empresa MKT – que produz a tecnologia de Confocal Laser Endomicroscopy –, o que enriqueceu a discussão aqui apresentada. Do Departamento de Física da Faculdade de Ciências da Universidade de Lisboa, fui orientada pelo Prof. Nuno Matela.

Palavras-chave: *Confocal Laser Endomiscroscopy*, Doenças Intersticiais Pulmonares, Análise Morfométrica, Análise de Textura, Fibras de Elastina.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

List of Acronyms

ADA Alveolar Ducts Area

ALAT Latin American Thoracic Association

AMC Academic Medical Center

ASM Angular Second Moment

ATS American Thoracic Society

BAL Bronchoalveolar lavage

BOW Bag-of-visual-words

CE European Conformity

CLE Confocal Laser Endomicroscopy

COPD Chronic Obstructive Pulmonary Disease

CT Computed Tomography

CTDs Connective Tissue Diseases

DAD Diffuse Alveolar Damage

DIP Desquamative Interstitial Pneumonia

DPLD Diffuse Parenchymal Lung Diseases

ERS European Respiratory Society

EVA Endomicroscopy Virtual Assistant

EVG Elastica van Gieson

FCFM Fibred Confocal Fluorescence Microscopy

FDA Food and Drug Administration

FFT Fast Fourier Transform

FOV Field-Of-View

GGO Ground-glass Opacities

GLCM Gray-Level Co-Occurrence Matrix

HE Hematoxylin-Eosin

HRCT High-Resolution Computerized Tomography

IDM Inverse Difference Moment

IIP Idiopathic Interstitial Pneumonias

ILD Interstitial Lung Disease

IPF Idiopathic Pulmonary Fibrosis

IPPFE Idiopathic Pleuroparenchymal Fibroelastosis

JRS Japanese Respiratory Society

LBP	Local Binary Patterns
LDAD	Longest Diameter Alveolar Ducts
LIP	Lymphocytic Interstitial Pneumonitis
LSU	Laser Scanning Unit
MDD	Multidisciplinary Discussion
MDT	Multidisciplinary Team
NSIP	Non-Specific Interstitial Pneumonia
OCT	Optical Coherence Tomography
OP	Organizing Pneumonia
PAP	Pulmonary Alveolar Proteinosis
pCLE	Probe-Based Confocal Laser Endomicroscopy
ROC	Receiver Operating Characteristic
RLL	Right Lower Lobe
SDAD	Short Axis Alveolar Ducts
SIFT	Scale Invariant Feature Transform
SVM	Support Vector Machine
TBLB	Transbronchial Lung Biopsy
TBLC	Transbronchial Lung Cryobiopsy
UIP	Usual Interstitial Pneumonia
VATS	Video-Assisted Thoracoscopic Surgery

1 INTRODUCTION

Identifying the type of Interstitial Lung Disease (ILD) a patient suffers from is a difficult process. Multidisciplinary discussion (MDD) is currently the gold standard to diagnose ILD patients: several specialist doctors compose a multidisciplinary team (MDT) that will consider the available clinical, radiologic and pathologic data from the patient to reach a conclusion.

Regardless of the effort and expertise associated with a multidisciplinary discussion, 10% of the patients are categorized as unclassifiable because of inadequate data or discrepancy between the existing data. The major cause for unclassifiable ILD is the absence of histopathological data associated with the risks of surgical lung biopsies [1][2].

It is very important to determine the patient's specific ILD due to its implications to his treatment and management [3][4]. It is particularly critical do distinguish the Idiopathic Pulmonary Fibrosis (IPF) patients from the non-IPF patients, since there are anti-fibrotic therapies – such as *Pirfenidone* – indicated for IPF which are very expensive, requiring certainty before being prescribed [5]. Also, treatment with immunosuppressive agents can work with the non-IPF, but it increases death and hospitalization among IPF patients [6][7].

The multidisciplinary discussion can benefit from the additional information offered by the Confocal Laser Endomicroscopy (CLE), an imaging technique which makes it possible to visualize the lung alveoli with micrometer resolution in a minimally invasive way, through a bronchoscopy. Through CLE, the elastic fibers networks in the alveoli are visible. There is evidence that the amount of elastin fibers increases, and the architecture of these fibers is altered in lung fibrosis. Lung Fibrosis is associated with some ILD including Idiopathic Pulmonary Fibrosis, being relevant identifying this condition.

So far, CLE movies are mostly being analyzed visually, and little objective and consistent information has been taken from these in ILD patients. It is possible to obtain more relevant data from the CLE movies by turning them into frames, pre-process them and extracting numerical features.

In this project, ILD patients had their lung alveoli imaged with CLE. After selecting the CLE movies and converting them into frames, the image quality will be improved, and the image analysis will be applied. Also, images with wider a field-of-view – mosaics – will be obtained. The image analysis will extract numerical values reflecting the state of the alveolar space, mainly the elastic fibers. The main goal of the project is to improve the CLE technique and increase its usability so that in the future, it can contribute to facilitate the stratification of ILD patients and eventually reduce the number of lung biopsies in these patients.

The mentioned project was developed through a 6-month internship at the department of Respiratory Medicine from the Academic Medical Center in Amsterdam, The Netherlands. At the Academic Medical Center (AMC), I was followed by the PhD students Lizzy Wijmans – medical doctor – and Paul Brinkman – biomedical engineer –, and supervised by Dr. Jouke Annema, MD, PhD, Professor of Pulmonary Endoscopy. This research group is focused on innovative techniques to image the pulmonary system. We had the opportunity to meet with the MKT company – manufacturer of the Confocal Laser Endomicroscopy technology – which enriched the discussion presented here. From the Physics department at Faculdade de Ciências da Universidade de Lisboa, I was supervised by Prof. Dr. Nuno Matela.

The resultant thesis is structured as follows: the Background gives an overview of the Interstitial Lung Diseases, their current diagnosing process, and presents the Confocal Laser Endomicroscopy

technology, ending with the state-of-the-art of lung alveoli imaging with CLE; the Materials and Methods section describes the clinical trial behind this study, the bronchoscopic procedure and the CLE data handling process – data selection, pre-processing and image analysis – including an explanation of how the gray level co-occurrence matrix is computed and how the texture features are extracted from it; the Results presents the patient group and some remarks from visually analyzing the CLE movies. Then, one patient is used as an example to represent the complete data handling process followed by an overview of those results for the entire data set; the Discussion comprises the strengths and issues associated with this project; finally, the last chapter summarizes the general Conclusions and Future Work.

2 BACKGROUND

This chapter covers an introduction to the Interstitial Lung Diseases (ILD) and the Confocal Laser Endomicroscopy (CLE) system, ending with the state-of-the-art of imaging the alveoli – alveoloscopy – with CLE. The ILD sub-chapter includes an overview of this broad and heterogeneous disease group, a presentation of lung fibrosis as an important element in some ILDs, and the description of the current diagnosing process including the role of radiological and histopathological data. Afterwards, the CLE system will be described, particularly its application in pulmonology, mentioning which structures from the lung are visible using this technique. Finally, the state-of-the-art will list some approaches from different research groups to obtain information from CLE images.

2.1 Interstitial Lung Diseases

Interstitial Lung Diseases (ILD) is a term which includes more than 200 lung disorders that share similar clinical, radiographic, physiologic, or pathologic manifestations. ILD are also named Diffuse Parenchymal Lung Diseases (DPLD) because they affect the lung parenchyma, including the alveoli, alveolar epithelium, capillary endothelium and in-between spaces, and also the perivascular and lymphatic tissues [8][6].

According to the American Thoracic Society (ATS) and the European Respiratory Society (ERS), the ILD can be classified into four main groups, as shown in Figure 2.1: 1) ILD of known-cause (e.g. drug side effects or occupational causes) or association (e.g. collagen vascular disease), 2) idiopathic interstitial pneumonias (IIP), 3) Granulomatous ILD (e.g. sarcoidosis) and 4) other forms of ILD (e.g. pulmonary alveolar proteinosis). The IIP (idiopathic interstitial pneumonias) have unknown cause (idiopathic) and are subdivided into idiopathic pulmonary fibrosis (IPF) and non-IPF IIPs [1].

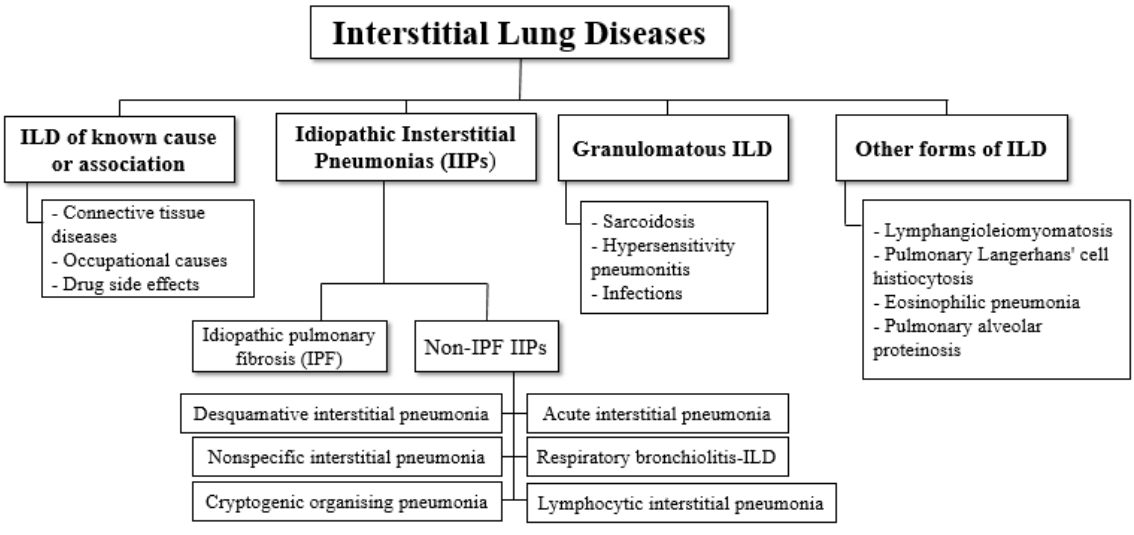

Figure 2.1 – Classification of interstitial lung diseases. Adapted from [1]

Among the ILD with known cause, the ones associated with occupational and environmental exposures are the most common, whereas in the remaining ILD with unknown cause, the most common diseases are sarcoidosis, idiopathic pulmonary fibrosis (IPF) [8].

In fact, IPF is one of the most prevalent ILD, with estimated incidence between 4 to 16 patients per 100,000 person-years [9]–[11]. This is a chronic and progressive fibrosing idiopathic interstitial

pneumonia (IIP), and the outcome of these patients is considerably worse comparing to others, which makes the distinction between IPF and other ILD more relevant. To identify IPF it is necessary to recognize the Usual Interstitial Pneumonia (UIP) pattern in radiography or histopathology [12][13].

2.1.1 Fibrosis in ILD

Inflammation and fibrosis are two important elements which are present in ILD in different degrees according to the specific disease. While inflammation is reflected by the increase in inflammatory cells and presence of nodules or edema, lung fibrosis is reflected by the additional collagen and elastin fibers.

The evaluation of the disease activity/state in ILD is more related to inflammation, requiring the assessment of number, type and state of activation of the inflammatory cells. In turn, the prognosis and the response to therapy is reflected by the extent of fibrosis and cellularity: predominant inflammation is associated with better prognosis and response to treatment, while predominant fibrosis indicates advanced stage and poor prognosis [6] [14]–[17].

Lung fibrosis is the scarring of lung tissue due to excessive deposition of extracellular proteins resulting in additional and disordered collagen and elastin fibers. Fibrosis results from chronic inflammatory reactions which can be triggered by persistent infections, tissue injury, autoimmune or allergic reactions, and exposures to chemicals or radiation.

The production of extra abnormal fibers results in architectural remodeling of the alveolar structures, tendency to collapse and loss of elasticity and lung function. Indeed, collagen and elastin fibers are the main components of the connective tissue network of the lung and their function is to provide elasticity and tensile strength. Whereas collagen fibers limit the extension of the alveolar ducts and sacs, the elastin fibers support their retraction. The disordered manner in which the new collagen and elastin fibers are produced leads to a "stiff" lung with reduced compliance, and impaired mechanical properties: the enhanced elastic recoil makes it harder to breath, specifically during early inspiratory phase [18] [19].

The production of tropoelastin, that generates the elastin, should take place merely during the human development period. However, in patients with pulmonary fibrosis, there is an increased tropoelastin expression in adulthood which results in the excessive elastin fibers in their lungs. Simultaneously, there is an increased production of collagen fibers by the fibroblasts and myofibroblasts which can form the so-called fibroblast foci [18][20].

2.1.2 Diagnosing ILD

The identification of the specific ILD is a difficult process due to the heterogeneity and variety of entities involved. Therefore, the multidisciplinary discussion (MDD) was introduced in the guidelines and is currently the gold standard to diagnose ILD patients [1]. In a MDD, several specialist doctors compose a multidisciplinary team (MDT) that will consider the available clinical, radiologic and pathologic data from the patient to reach a conclusion.

The MDT must have a pulmonologist, radiologist, and a pathologist. At the Academic Medical Center in Amsterdam (AMC), this team also includes specialists in pulmonary endoscopy, surgery, radiotherapy and oncology.

The diagnostic process of a suspected ILD starts before the MDD with the evaluation of the history, physical examination, chest radiographs and lung function tests from the patient. Regarding the patient's history, the environmental and occupational exposures, smoking status and drug use are important details. At this point, the clinician may consider the disease to be an ILD other than IIP, or an IIP. In this case, a high-resolution computerized tomography (HRCT) scan is indicated, which normally has four outcomes:

- Distinctive features of Usual Interstitial Pneumonia (UIP) (associated with IPF)
- Atypical clinical or radiological features for IPF
- Distinctive features of another ILD
- Suspected features of "other forms of ILD" (see Figure 2.1)

The main goal of the HRCT is to identify the patients with distinctive IPF features. It also gives information about the extent of the abnormalities on both lungs. Patients with atypical or indistinctive features may be subject to bronchoalveolar lavage (BAL) and collection of tissue samples through one of these procedures: transbronchial lung biopsy (TBLB), transbronchial lung cryobiopsy (TBLC), or a surgical lung biopsy, namely a video-assisted thoracoscopic surgery (VATS).

BAL is a procedure which uses a flexible bronchoscope to pump a saline solution into the bronchial tree and then collect it. The cytological description of the BAL fluid gives information about the predominant inflammatory cellular pattern (increased lymphocytes, eosinophils, or neutrophils). This information is important particularly when there isn't a confident UIP pattern and, for some cases, avoids more invasive procedures [21]–[23].

When there is a need to take biopsies from the patient, although VATS can provide bigger samples that lead to more confident diagnosis, bronchoscopic biopsies are safer for the patient, avoiding a surgical procedure. In fact, most ILD patients have multiple comorbidities and cannot undergo such an invasive procedure. VATS has a postoperative mortality rate between 2% and 4.5% and requires more financial costs. The bronchoscopic procedures, TBLB and TBLC, are safer than VATS but are not harmless: death is infrequent, but there are risks involved such as pneumothorax and bleeding. Comparing TBLB and TBLC, the first is more frequently associated with nondiagnostic specimens due to their small size and "crush" artifacts, while cryobiopsies are usually well preserved and the bigger area of the specimens leads to a higher diagnostic yield [24]–[27]. Both TBLB and TBLC are usually guided using fluoroscopy, through a C-arm x-ray system, which provides a 2D image. The resolution of the fluoroscopic system is limited, and the radiation dose involved is harmful for the patient and the medical team [24].

Afterwards, at the multidisciplinary discussion, the patient's case is presented including all the available data such as the HRCT images, cytological description from the BAL (bronchoalveolar lavage) and pathological description from the biopsies.

Regardless of the effort and expertise associated with a multidisciplinary discussion (MDD), 10% of the patients are categorized as unclassifiable. The major cause for unclassifiable ILD is the absence of histopathological data associated with the risks of surgical lung biopsies, but other reasons include inadequate clinical, radiologic, or histopathological data, discrepancy between the data, nondiagnostic biopsies and discrepancy between biopsies from different lobes [1][2].

It is very important to determine the patient's specific ILD due to its consequences to the treatment and management [3][4]. It is particularly critical do distinguish the Idiopathic Pulmonary Fibrosis (IPF) patients from the non-IPF patients, mostly because of the anti-fibrotic treatment recommended for IPF patients, such as the medication named *Pirfenidone* [5]. This medication is proven to reduce disease progression but is associated with side effects (the most common are gastro-intestinal and skin-related adverse events)[28]. The fact that treatment with *Pirfenidone* is very expensive – costing the Dutch government, for example, approximately 14 million euros per year (approximately 5000 euros per patient per year) – demands more certainty in the IPF diagnosis before prescribing that medication. Conversely, treatment with immunosuppressive agents can work with the non-IPF patients, but it increases death and hospitalization among IPF patients[6][7].

2.1.2.1 Radiography in the diagnosis of ILD

In the CT images, ILD are characterized by infiltrative opacification in the periphery of the lung. As the alternative designation for ILD – diffuse parenchymal lung diseases – suggests, they are diffuse/spread.

Figure 2.2 presents 2 HRCT scans showing a pair of healthy lungs and the lungs from an ILD patient. It is evident that the images are very different. It can be seen that the diseased lungs have thickened alveoli septa.

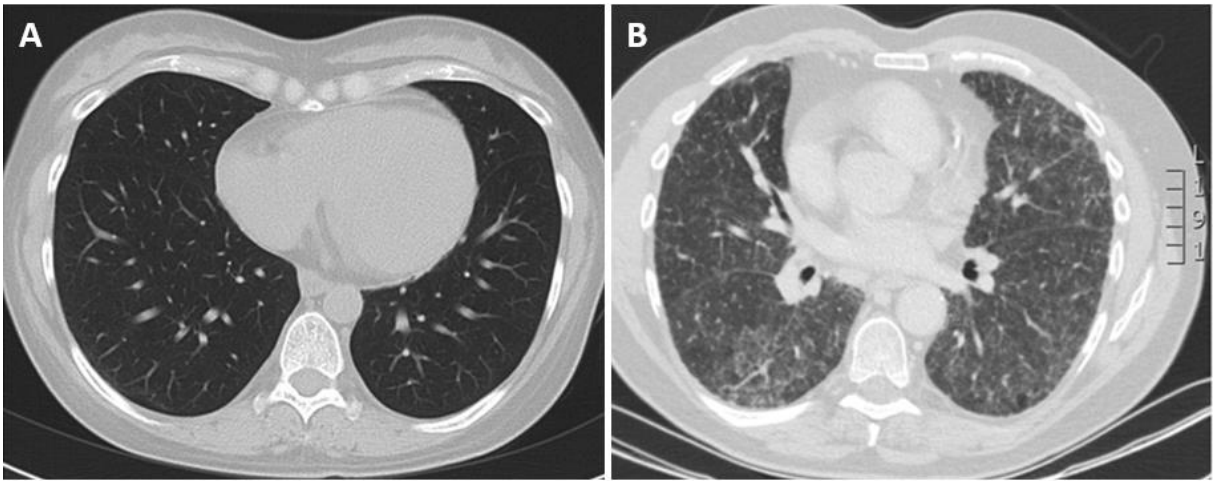

Figure 2.2 – Comparison between radiography of A) normal lungs (HRCT) and B) diseased lungs from an ILD patient (CT-scan). Image rights: L. Wijmans, Academic Medical Center, Amsterdam, Netherlands

According to K.O. Leslie [29], there are 4 different radiological patterns which can be identified in the HRCT of an ILD patient (Table 2.1):

Table 2.1 – Radiologic Patterns of Interstitial Lung Diseases identified in HRCT images according to K.O. Leslie [29]

Radiologic Pattern	Examples of ILDs associated with this pattern
"Ground-glass opacity and consolidation", increased attenuation (edema and infection)	Hypersensitivity Pneumonitis, Idiopathic Interstitial Pneumonias (NSIP, LIP and DIP), Pulmonary Alveolar Proteinosis, fumes and toxin injury
"Fibrosis", reticulation with parenchymal distortion	IPF, Sarcoidosis, LIP, advanced Pulmonary Langerhans' cell histiocytosis
"Nodules" (can be large or small, singular or multiple)	Hypersensitivity pneumonitis, Pulmonary Langerhans' cell histiocytosis, Sarcoidosis, Silicosis and silicate disease (also carcinomas and sarcomas)
"Mosaic patterns and cysts"	Langerhans cell histiocytosis, Small-airways disease with constrictive bronchiolitis, Lymphangioleiomyomatosis

"Ground-glass opacities (GGO) and consolidation" appear as increased attenuation in the radiographic image. While in GGO the "attenuation" is not dense enough to obscure the bronchovascular markings, when "consolidation" is referred, it means those markings are not visible. Nodules are also areas of increased attenuation normally with discrete borders and sizes between a few millimetres up to 3 cm.

When lung fibrosis is identified, it is usual to assess its degree and extent, and the presence or absence of “honeycombing” which is associated with loss of architecture and appears as decreased attenuation. Honeycomb cysts are large, with a uniform size have thick and well-defined walls (thickness between 3 and 10 mm). These cysts are often stacked in one or more layers. The term “reticulation” refers to a netlike pattern which results mainly of thickened inter- and intralobular septa, connecting with one another at different angles.

The cysts referred in the last radiological pattern present in Table 2.1 are also areas of low attenuation with round shape having a thin wall less than 2 mm thick.

Figure 2.3 shows some examples of radiological images (axial/coronal HRCT or contrast enhanced chest CT) from ILD patients’ lungs. The caption includes the corresponding patient diagnosis and mentions some of the distinctive features present. As the figure illustrates, this disease is very heterogeneous with each entity showing different radiological characteristics, which at the same time are similar in some way. Also, the disease state will influence the severity of these findings, making it even harder to reach a conclusive diagnosis.

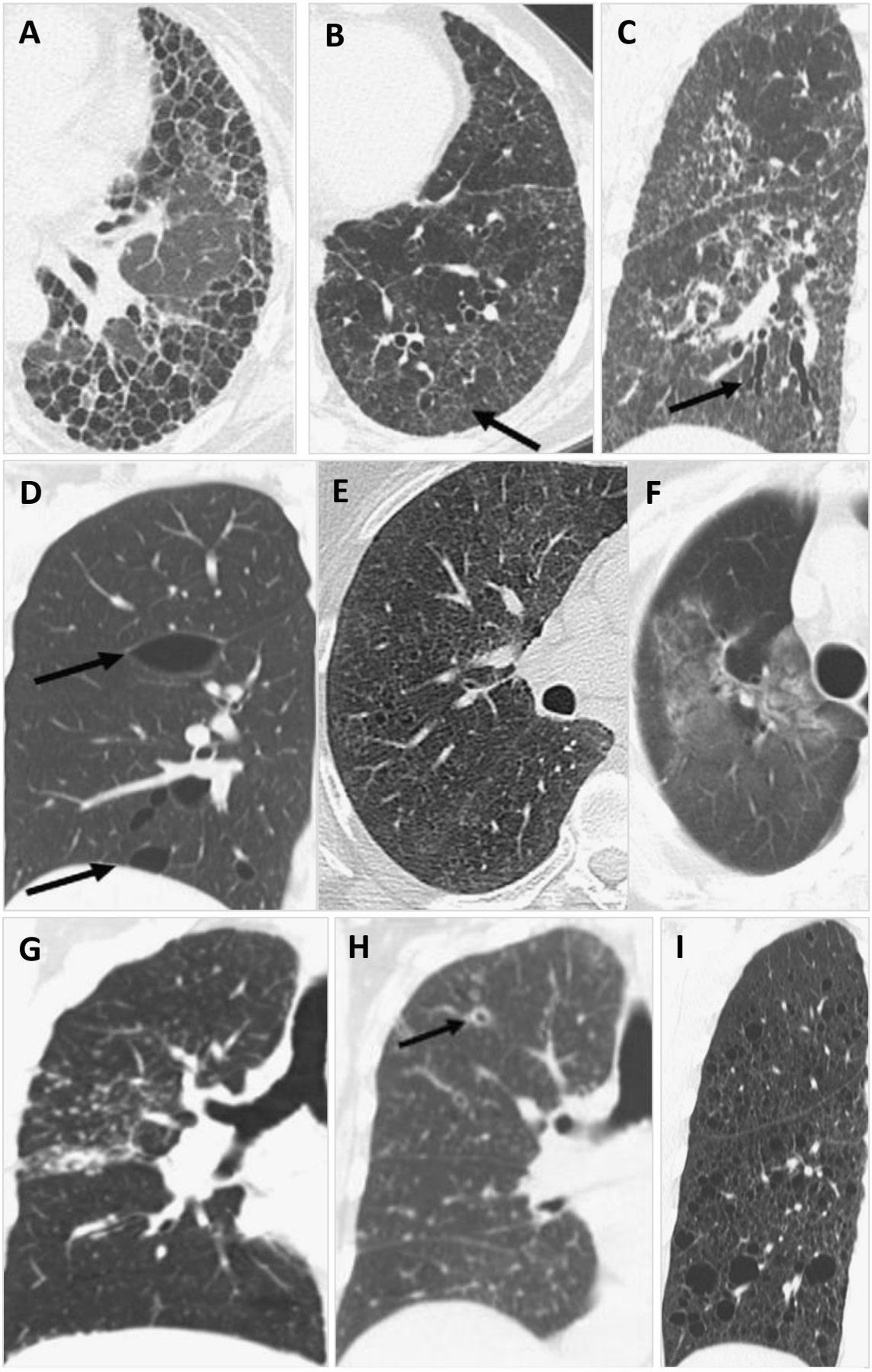

Figure 2.3 – caption on the next page

Figure 2.3 – (on the previous page) Some examples of radiographic images (CT or HRCT) of ILD patient's lungs. Images were adapted from [29]

A) Idiopathic Pulmonary Fibrosis (IPF) – stacked thin-walled cysts; fibrosis is most severe in the basilar and peripheral portions of the lungs.
B) Nonspecific Interstitial Pneumonia (NSIP) – basilar predominant reticulation (arrow).
C) Fibrotic NSIP – basilar predominant ground-glass opacity, reticulation, and traction bronchiectasis (arrows).
D) Lymphocytic Interstitial Pneumonitis (LIP) – subpleural and bronchovascular thin-walled lung cysts (arrows).
E) Desquamative interstitial pneumonitis (DIP) – diffuse ground-glass opacity with superimposed small cysts.
F) Cryptogenic Organizing Pneumonia (COP) – central ground-glass opacity.
G) Sarcoidosis – multiple small nodules primarily in the midlung zone.
H) Pulmonary Langerhans Cell histiocytosis (LCH) – small mildly thick-walled lung cysts (arrow) and multiple centrilobular nodules.
I) Lymphangioleiomyomatosis (LAM) – diffuse thin-walled lung cysts

2.1.2.2 Histopathology in the diagnosis of ILD

In histopathology, the usual procedure of ILD diagnosis includes staining the biopsy samples with hematoxylin-eosin (HE). The pathologist will mainly search for the presence of dense collagen deposition and fibroblastic foci in order to assess lung fibrosis and its distribution. Also, he will look for signs of inflammation such as cells, infiltrates, edema, and nodules [18].

As mentioned, the biopsies can derive from a bronchoscopic (TBLB or TBLC) or surgical (VATS) procedure. From the tissue samples obtained, six histopathological patterns can be found, according to K.O. Leslie [29] (Table 2.2):

Table 2.2 – Histological patterns identified in biopsies from ILD patients according to K.O. Leslie [29]

Histological Pattern	Basic elements of the pattern
Acute lung injury	Interstitial edema, intra-alveolar fibrin, reactive type 2 cells
Fibrosis	Dense collagen deposition in the lung parenchyma, some degree of structural remodeling with alveolar loss
Chronic inflammatory (cellular) infiltrates	Chronic inflammatory cells with diffuse distribution within alveolar walls
Alveolar filling	Alveoli filled with cells or non-cellular material
Nodules	Nodules that can be well or poorly formed, large or small, single or numerous.

Note that the “Fibrosis” and “Nodules” patterns are common between radiology and histopathology.

The UIP (usual interstitial pneumonia) pattern which identifies the IPF (idiopathic pulmonary fibrosis) clinical syndrome is included in the Fibrosis pattern both in radiology (Table 2.1) and histopathology (Table 2.2). The HRCT features which indicate UIP are: reticular opacities, clustered cystic airspaces – honeycombing –, and ground glass opacities. The histopathology features relative to UIP are: fibrotic areas alternate with normal (or almost normal) areas – heterogeneity – when seen in low magnification, and mild inflammation with interstitial infiltrate of lymphocytes and plasma cells. The fibrotic areas consist of dense collagen deposition and may have fibroblast foci (foci of proliferating fibroblasts and myofibroblasts) and honeycombing (cystic airspaces that can contain mucus and inflammatory cells) [30][13].

In summary, the HRCT imaging is not always enough to make the ILD diagnosis given its resolution and field-of-view (FOV), and histopathology is not always an option because of the associated risks. Even in the presence of histological material, it is not certain that a conclusion will be reached due to the possibility of bad quality or nondiagnostic tissue samples. Therefore, there is a need for an alternative/complementary diagnostic technique, with a better resolution than HRCT and less invasive in comparison with histopathology.

2.2 Elastin Fibers in the ILD Diagnosis

The collagen and elastin fibers present in the lung are intermingled and form a continuum which connects the alveolar ducts, alveolar septa, blood vessels and pleura. The fibers are more dense in the alveolar mouths and become thinner in the alveolar septa [19]. Figure 2.5 and Figure 2.4 A) show the elastin fiber network from a healthy human lung, where the respiratory ducts are the large channels that narrow into alveolar ducts leading to the alveolar sacs which contain multiple alveoli. The blood vessels can have a single layer or multiple layers (in thick vessels) of elastin fibers, and the outermost fibers are continuous with the alveoli fibers. The pleura is the membrane which covers the lung and is visible in Figure 2.5 and Figure 2.4B as a dense matrix of fibers.

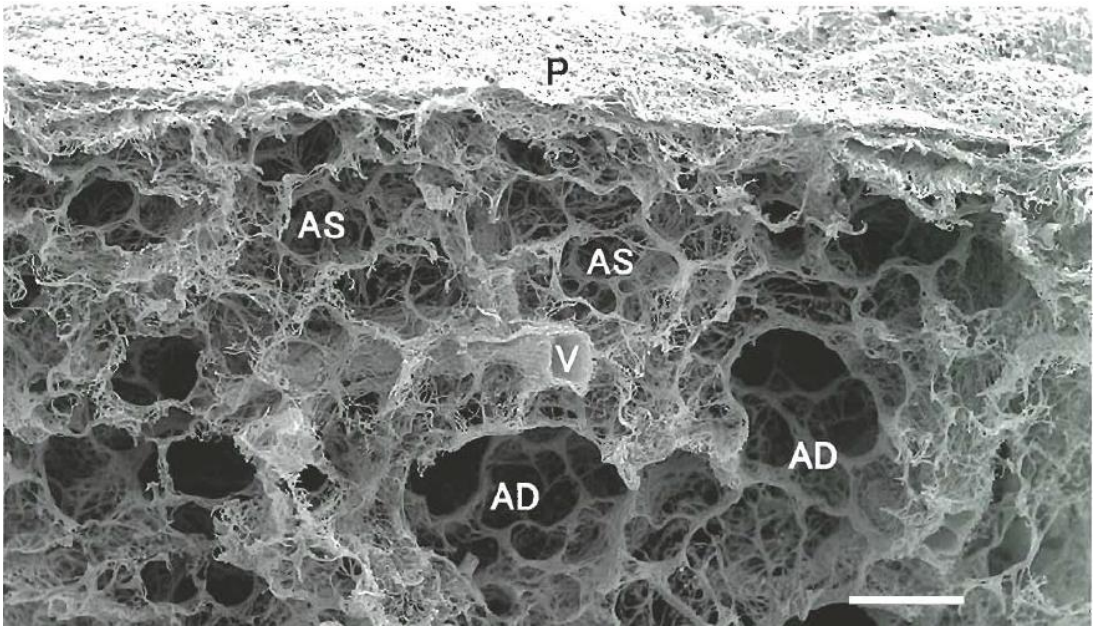

Figure 2.5 – Elastin fiber networks in the human lung. Scanning electron micrograph with scale bar = 200µm. Legend: AD- Alveolar duct, AS-Alveolar sac, P-Pleura, V-Blood vessel. Adapted from [19]

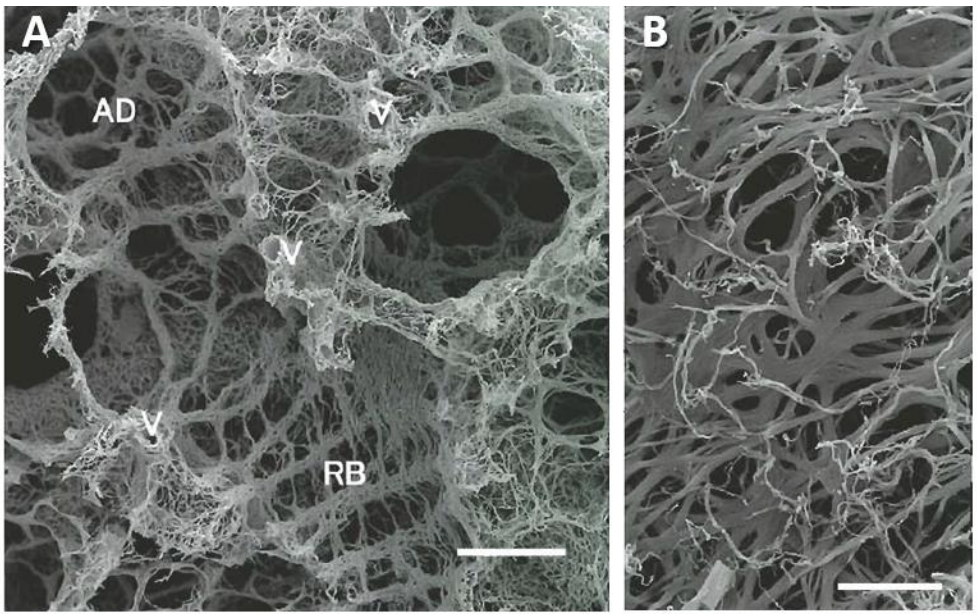

Figure 2.4 – A) Elastin fiber networks in the respiratory bronchiole in the human lung. Scanning electron micrograph with scale bar = 200µm. Legend: AD-Alveolar duct, RB-Respiratory bronchiole, V-Vessel. B) Elastin fiber networks in the outer surface of the pleura in the human lung. Scanning electron micrograph with scale bar = 10µm. Adapted from [19]

As referred in section 2.1.1, the fibrotic process in ILD consists in the production of extra collagen and elastin fibers. However, the usual procedure done by the pathologist to recognize lung fibrosis and assess its distribution is to search for the dense collagen deposition and associated fibroblastic foci in the samples stained with hematoxylin-eosin (HE). The identification of elastin to evaluate fibrosis in ILD is unusual and not mentioned in the guidelines. This can be justified by the fact that HE is one of the most common staining techniques in histology and it makes possible to identify collagen. However elastic fibers stain poorly with eosin being difficult to recognize in HE preparations [18][13][31].

Elastica van Gieson (EVG) is a stain containing picric acid and acid fuchsin which differentiates the connective tissue components: collagen appears in red and elastin in black [32]. Some research groups have been making quantitative analysis of the elastin fibers in lung biopsies in ILD using EVG or other elastic stains. As a result, some studies have shown that the deposition of elastic fibers – elastosis – parallels the collagen deposition in idiopathic interstitial pneumonias (IIP). The quantitative analysis of elastic fibers in tissue samples from IIP patients showed that the proportion of elastin fibers was higher in IIP than in normal lungs. A study which included IIP patients with 4 different histological patterns – DAD (diffuse alveolar damage), OP (organizing pneumonia), NSIP (non-specific interstitial pneumonia) and UIP (usual interstitial pneumonia) – concluded that, although the increase in elastin fibers compared to normal lungs, there was no significant differences between the 4 patterns in terms of elastic fibers amount [33]. In turn, another study which compared only two IIP – IPPFE (idiopathic pleuroparenchymal fibroelastosis) and IPF (which has the UIP histological pattern) – found out more than twice the amount of elastin fibers in IPPFE than in IPF lungs [34].

Regarding collagen, the UIP pattern reveals a significantly higher proportion of collagen fibers comparing to DAD, OP, NSIP, and of course normal lungs. Also, while DAD, OP, NSIP show similar increase in the collagen and elastic fibers, UIP reveals a higher collagen deposition relative to the elastin deposition.

Although the clinical relevance of the increased elastin fibers in ILD is not fully studied, it was found to be correlated to the extent of lung fibrosis and severity on HRCT and histopathology. Furthermore, the augmented proportion of elastin fibers in histology specimens is an independent predictor of a poor prognosis in IPF patients: patients with augmented elastic fibers have reduced lung function, and worse survival.

Besides the amount of fibers, if their appearance is evaluated, one observes a distorted elastic network in the areas of lung remodeling and the shape of the new elastic fibers appears “frayed”, just like the fabric fibers from an old piece of clothing.

In conclusion, in IIP an increased amount of elastin fibers is found and their shape is altered [33]–[36][14].

It is also important to be familiar with the appearance of the pleura, because it can avoid a pneumothorax during lung biopsies. If a pneumothorax occurs, the biopsy will certainly include pleura, thus, if that tissue is imaged *ex vivo*, the increased amount of elastin fibers characteristic of pleura should not be mistakenly described as alveolar fibrosis.

2.3 CLE System

Confocal Laser Endomicroscopy (CLE) is an imaging technique which allows for real-time microscopic information *in vivo* through an endoscopic procedure in a minimally invasive way. It also allows to image *ex vivo* after a biopsy is collected. CLE emits low-power blue laser light, which is reflected from the tissue and refocused on the detection system by the same lens. The term "confocal" means that the illumination and collection systems are in the same focal plane. The existing pinhole ensures that only the light coming from the focus plane is collected and light from out of focus planes is excluded, as observed in Figure 2.6, which guarantees a spatial resolution allowing cellular imaging. CLE takes advantage of natural fluorescence of the tissues being examined, but topical or intravenous contrast agents can also be used to enhance contrast, although it is not common in pulmonary imaging.

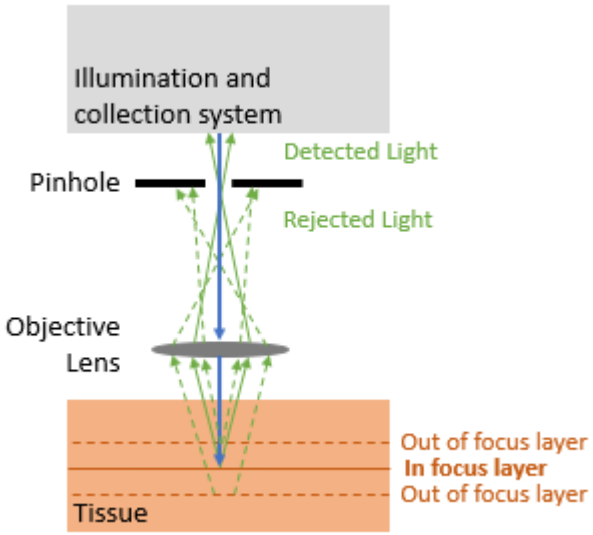

Figure 2.6 – Confocal system in CLE diagram. Adapted from [38] [39]

Cellvizio from Mauna Kea Technologies (in Paris, France) produces the main CLE systems for application on pulmonology certified by CE and FDA which are commercially available. This technology is sometimes designated by fibred confocal fluorescence microscopy (FCFM). The designations endoscopic confocal fluorescence microscopy, fiberoptic confocal system and high resolution microendoscopy are less common.

2.3.1 CLE System in Pulmonology

In standard bronchoscopies, it's not possible to reach the distal lung, so the real-time images won't go beyond second or third order bronchi. Conversely, if you reach the distal lung with a biopsy probe, the pathologist will examine the collected tissue through a two-dimensional sample devoid of air and the original three-dimensional structure of the alveoli. Those gaps intent to be filled by the probe-based CLE (pCLE) from Cellvizio: the pCLE probe is inserted in the working channel of a bronchoscope and extended further into the terminal bronchi or alveoli. It makes it possible to do an "alveoloscopy" with microscopic resolution, through a grayscale video sequence, while the patient is breathing. This system has a fixed confocaldepth (0 to 50 μm) meaning it scans in a single plane and can't image at different depths.

Table 2.3 – Cellvizio AlveoFlex confocal miniprobe characteristics [37]

	Diameter	Compatible operating channel	Length	Maximum # of uses	Field of view	Resolution	Confocal depth
AlveoFlex	1.4 mm	≥ 1.9 mm	3 m	20	Ø 600 µm	3.5 µm	0-50 µm

The pCLE miniprobe from Cellvizio used in bronchoscopy is called AlveoFlex and its characteristics are summarized in Table 2.3. This miniprobe is a bundle of 30.000 optical fibers. Each fiber has a 1.9 µm diameter and is a "point detector" with miniaturized optics on its distal tip that focuses the light into the tissue, collects the fluorescent signal and acts as a pinhole (as represented in Figure 2.6). The non-imaging space between the fiber cores can limit resolution, but it is compensated by the image processing algorithm which interpolates the information from the surrounding fibers. The AlveoFlex has a metal-containing material on the tip, so it can be tracked in the CT scan. The comparison between the field-of-view (FOV) obtained in CLE and CT (and OCT) is in Figure 2.7, showing the 0.6mm FOV for the pCLE. There are other Cellvizio miniprobes which differ in their distal optics and bundle characteristics (flexibility, size, focal depth, field of view, sensitivity and resolution) and applied to different indications such as urology, gastric diseases, biliary and pancreatic structures, Barrett's esophagus, pancreatic cysts and colorectal lesions.

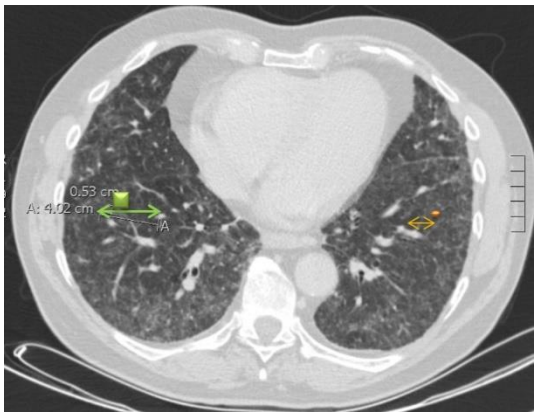

Figure 2.7 – Comparing the FOV in CT with pCLE. pCLE is represented by the orange dot, FOV=0.6mm. The green dot represents another technique, OCT with FOV=0.5mm. The arrows represent the pull-back distance / length of trajectory which is approximately 1cm for pCLE and 4cm for OCT. Image rights: L. Wijmans, Academic Medical Center, Amsterdam, Netherlands.

The Cellvizio system (see Figure 2.8) contains the laser scanning unit (LSU), the image processing software and the confocal miniprobe. The LSU contains the blue laser diode which beam has a 488 nm excitation wavelength that will travel by a system of lenses and mirrors thinning it to a micron scale diameter until it reaches the miniprobe. Then, in the AlveoFlex, the beam travels inside one of the 30000 optical fibers. The scanning process happens in high speed with each fiber being injected 12 times per second: a 4 kHz oscillating mirror makes the horizontal line scanning and a galvanometric mirror makes the frame scanning. This allows for real-time imaging with an output grayscale video with 9 frames per second. When the blue laser light reaches the patient's lung tissue, it will scan a circular area with a 600 µm diameter. The natural occurring fluorescence of the lung will be collected by the miniprobe and guided back to the LSU, with a lateral resolution of 3.5 µm. After the images are processed in the image

processing software, the output is a grayscale video that can be watched on screen in real-time and recorded.

Figure 2.8 – CLE system from Cellvizio constituted by the confocal miniprobe attached to the laser scanning unit (LSU), and the image processing software. The grayscale video images are visualized on screen. On the floor there is a foot pedal which makes it easier to start and stop recording sequences during the endoscopic procedure. Adapted from [74]

Besides interpolating the image data in the space between fibers, the image processing algorithm in the Cellvizio machine corrects for artifacts (occurring while the beam travels inside the fiber bundle) and adjusts the intensity of the image during the acquisition according to the collected intensity. This "laser automatic control" ensures that the structures in the images have good contrast but makes it impossible to make reliable comparisons between frames regarding intensity values. In the mkt file, there is no record of the information related to the intensity adjustments that were made. In this sense, it is not possible to revert the adjustments once the imaging procedure was done. Thus, to make intensity analysis, the laser automatic control must be deactivated by a technician from the MKT company before the endoscopic procedures take place.

The Cellvizio system has a built in Endomicroscopy Virtual Assistant (EVA) with an intuitive interface making it easier to fully use the potential of this technology. Some of EVA features are: "The Atlas", a database of reference videos for each CLE application (esophagus, colon, lung…), the "SmartReview" which highlights the video segments considered stable by the software and "FastExtract" that allows to immediately export the stable videos selected by the "SmartReview".

To use the Cellvizio system is intuitive and it only requires the clinician who performs the bronchoscopy to carry out brief training. Also, the patient does not need any additional preparation. The AlveoFlex probe is compatible with all bronchoscopes operating channels that currently exist and can be used in 20 procedures. During the bronchoscopy, the clinician knows where the bronchoscope is positioned using the standard white light images, and when it reaches its access limit, the AlveoFlex is pushed further into the alveoli. From this point, there is no certainty of the precise brunch where the probe is. Thus, the precision of the miniprobe's location is limited to the lung region associated to the position of the bronchoscope [38][39][40].

Cellvizio provides an imaging software called Cellvizio Viewer that reads the video sequences exported from the Cellvizio system (in format mkt) and allows to analyze them as a video or frame by frame. Besides making some "manual" measurements in the images such as distances or angles, the software has a feature named Mosaicing which "stiches" consecutive frames from a stable CLE video, making it possible to virtually increase the field of view.

2.3.1.1 Visible structures with CLE in Pulmonology

The 488 nm laser light emitted by the CLE makes it possible to image the bronchial wall and the alveoli due to the elastin present in these structures, with elastin being the main fluorescent signal detected. Some autofluorescent cells namely macrophages are also visible, as well as some lipids and other proteins. At this wavelength, the fluorescence yield of collagen is at least one order of magnitude smaller than elastin's, what justifies the fact that collagen is not visible in CLE images [41]–[44].

Figure 2.9 shows two frames from a CLE movie of healthy lung tissue available at the Cellvizio website [45]. The healthy alveoli are formed by regular alveolar septa, which elastin fibers are visible. The blood vessel present on the right-side image is easily identified because it contains multiple fibers. These images are comparable with those in section 2.2 where the elastin network of the lung is shown.

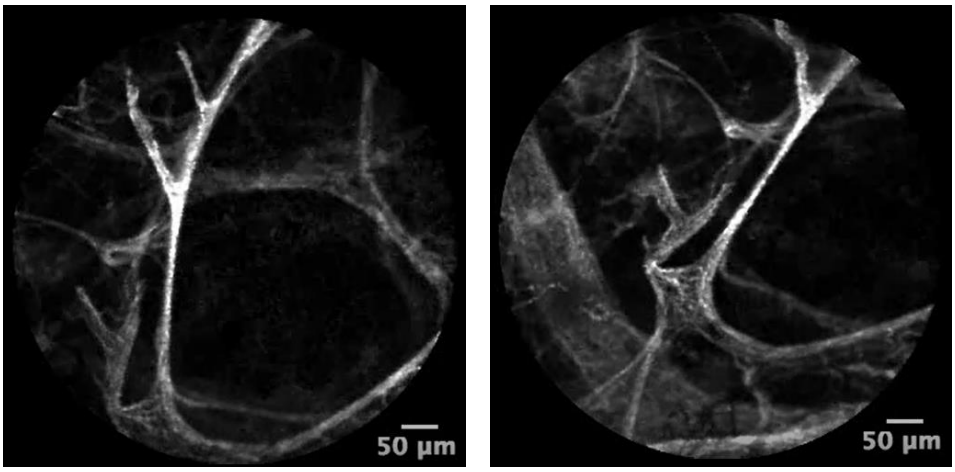

Figure 2.9 – pCLE images of healthy alveoli. The elastin fibers composing the round shaped alveoli are visible. The alveolar septa are thin. At the right-hand image, there is a blood vessel. These images were taken from the videos supplied in the training section from the Cellvizio website [46]

From studying smokers, it has been shown that tobacco tar is highly fluorescent at 488 nm excitation, acting as an exogenous fluorophore which enhances the signal detected by the CLE from the macrophages containing tobacco tar. Also, in patients with pulmonary alveolar proteinosis (PAP) - a rare type of ILD in which the surfactant accumulates in the terminal airways and alveoli - CLE images showed the presence of highly fluorescent globular and granular structures (with diameter around 100–300 mm), floating in a slightly fluorescent alveolar fluid. Those globular and granular structures are

lipoproteinaceous material typical of PAP, which can be differentiated from macrophages since the latter are smaller and have heterogeneous fluorescence (see Figure 2.10). Moreover, investigators said the examination of the BAL fluid with the CLE probe can be sufficient to diagnose PAP [43], [46]–[48].

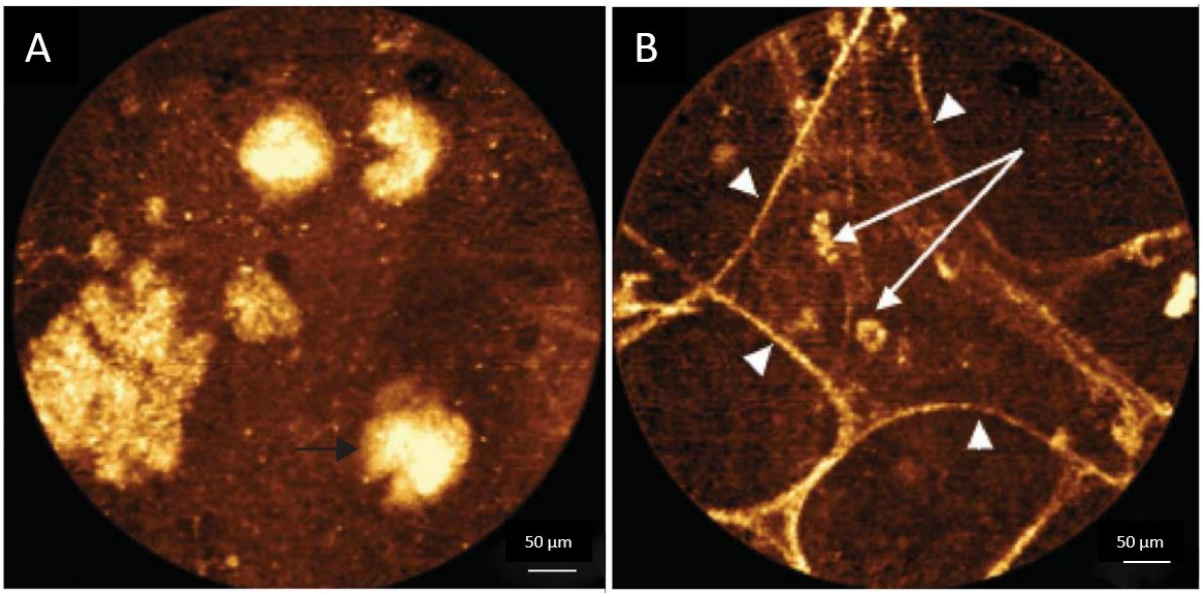

Figure 2.10 – CLE images from a patient with pulmonary alveolar proteinosis (PAP) showing A) lipoproteinaceous material typical of PAP: highly fluorescent globular structures (black arrow), with well-defined edges and homogeneous fluorescence; B) normal alveolar area with typical alveolar macrophages (white arrow) and normal alveolar elastin fibers (arrowheads). The alveolar macrophages are smaller than the lipoproteinaceous material and display heterogeneous fluorescence. Adapted from [46]

N. A. Bhatt and colleagues [49] imaged with CLE a sample of lung tissue containing pleura. They stated that pleura can be identified by its architecture, which is visibly different from lung parenchyma, airway mucosa and pulmonary lesions, showing a very organized cross-hatching of fluorescent linear structures. They suggest CLE should be used to guide biopsies and avoid pneumothoraxes.

As it is common in gastroenterology imaging, researchers questioned if exogenous contrast agents would improve the quality of CLE images in pulmonology. Fluorescein sodium could potentially affect the image in a positive way since it is a fluorescent contrast agent that absorbs blue light with peak excitation at wavelengths between 465-490 nm. However, it has been concluded that the administration of fluorescein sodium does not provide considerable additional information to the images since it ends up causing a general "white-out" of the image which does not contribute to the distinction of structures and thus to the diagnostic value [50][51].

Although the confocal microfibers in CLE allow a very high-resolution image from the focal plane at a 0-50 µm depth, as the probe is being pushed by the clinician, the more the alveoli will be compacted and more fibers from several structures will appear on the same plane.

2.4 State-of-the-Art of Alveoloscopy with CLE

Research articles regarding confocal laser endomicroscopy started to be published in 2004. From 2004 until 2008 there were no more than 10 publications per year on this subject, and most of them were introductions to the technology. Indeed, researchers were very enthusiastic about the imaging quality and the safety of the procedure. From 2004 up to now, more than 90% of CLE publications are gastroenterology related. Urology is another specialty of interest for this technique and there are some studies about microvasculature, nerve surgery and brain tumors. In 2009, the first article on CLE in pulmonology – "Confocal fluorescence endomicroscopy of the human airways" by Luc Thiberville [42] – was published, describing the applications and limitations of CLE for *in vivo* proximal and distal lung explorations.

In pulmonology, CLE has been applied mostly to the study of lung cancer, pulmonary nodules and ILD. Aspergillosis, chronic obstructive pulmonary disease (COPD), microlithiasis and acute lung injury are other conditions present in publications regarding CLE. There are also publications with a "non-focused" patient group, including the comparison of healthy with smokers, for example.

When CLE is applied to ILD patients, it is possible to split the methods present in research articles into 3 categories: those who make a visual description of the images captured by the pCLE probe (typically case reports of one disease) [46], [48], [52]–[54], those who make a qualitative analysis on the images by visual scoring [50][55] and those who attempt to make a quantitative analysis using more objective parameters to evaluate the images [56]. These 3 types of approach to the examination of the CLE images – visual description, qualitative analysis and quantitative analysis – will be discussed below. Afterwards, a brief description of a specific quantitative analysis – the texture analysis – used in CLE images from different clinical fields will be presented.

2.4.1 Visual Description of CLE images from ILD patients

A few studies have been published in which the main goal is to make a description of the features observed in the CLE images of ILD patients, whether they are focused in a specific ILD entity or cover general entities shared by several ILD diseases. Of course, most of the studies where qualitative and quantitative analysis is performed also include visual descriptions of great utility.

Some examples of visual descriptions will be presented hereafter:

- Patient with early pulmonary fibrosis (identified in CT) and a UIP pattern (in the histopathology): «pCLE imaging demonstrated the absence of the differentiation of alveolar walls resulting in their complete destruction with well distinguishable fibrillar branching fluorescent structures and no cells» [54]
- Patient with invasive pulmonary aspergillosis: «completely destroyed alveolar elastin network and well distinguishable fibrillar branching fluorescent structures in the lung parenchyma» [54]
- Patient with UIP pattern (in the histopathology): «the elastin structure of the alveoli is significantly deteriorated and disintegrated»[52]
- Patient with Pulmonary Alveolar Microlithiasis: «the alveolar opacities showed heterogeneous auto-fluorescence intensity with linear or round zones of decreased fluorescence. »[57]
- Patient with emphysema: «destruction and disarrangement of elastic fibers (…) and increased cross sectional area of alveolar openings in emphysema»[58]

Note the described elements were identified in some segments of the CLE movies, and might not be present in the entire sample (the same happens in histology).

2.4.2 Qualitative Analysis of CLE images from ILD patients

Other research groups studying CLE tried to identify and assess visual patterns in the images, and the presence of entities such as cells using grading scales.

O. Danilevskaya, et.al. [48] presented a 6-point score to assess the amount of floating intra-alveolar substances containing alveolar macrophages, with 0 meaning the absence of intra-alveolar complexes and 5 its most noticeable presence: 1 point – single complexes; 2 points – less than half of the FOV is covered; 3 points – half of the FOV is covered; 4 points – more than half of the FOV is covered; 5 points – the FOV is completely covered.

F. Reichenberger [55] evaluated the CLE images from patients with different ILDs according to the parameters: degree of distortion of the hexagonal alveolar pattern, density of alveolar structures and presence of "foamy" alveolar macrophages. These were visually assessed by independent investigators blind to the diagnosis using a scale from 0 to 3. It concluded, for example, that IPF patients presented severe distortion (level 3).

A. Wellikoff's colleagues [59] assessed CLE images from patients with lung cancer. They identified the presence (or absence) of the following features: fragmented elastin, friable elastin, black holes, elastin disorganization or clumping, cellular infiltrate, black holes along vessels, thickened septa and dense elastin.

R. Newton [50] searched for the presence of decreased autofluorescence of individual elastic structures, unrecognizable structures (and decreased distinctiveness), distorted wall morphology, autofluorescent macrophages and snapped septal walls or vessels.

In 2016, P. Meng and colleagues [56] analyzed the fiber patterns in CLE images, with the intention of correlating the results with CT findings and patients' final diagnosis. The group defined 6 patterns, shown in Figure 2.11.

A. Normal: the alveolar rings show normal regular elastin framework
B. Others/Non-specific: no regular elastin framework of the alveolar rings; loose fibers not recognized as part of alveoli or bronchial wall
C. Increased Fibers: increased number of fibers with normal alveolar structures still visible
D. Densely packed fibers: alveolar structures are absent
E. Hypercellular: cellular structures are visualized within the alveolar space
F. Thickened Fibers: thickened fibers with normal alveolar structures still visible

They found a positive association between this Hypercellular pattern and the "inflammatory pattern" they defined in CT images which included ground-glass opacification, nodular opacities, consolidation and bronchial wall thickening.

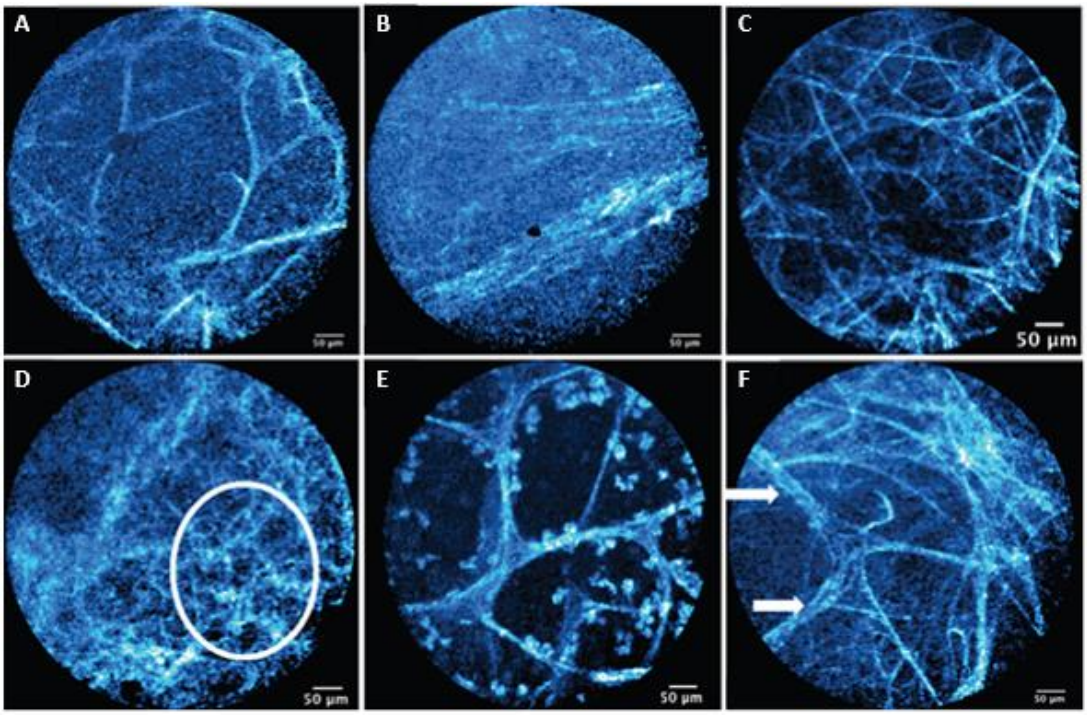

Figure 2.11 – CLE patterns in ILD patients defined by P. Meng. A) Normal, B) Others/Non-specific, C) Increased Fibers D) Densely packed fibers, E) Hypercellular and F) Thickened Fibers. Image from [56] *(adapted)*

2.4.3 Quantitative Analyses of CLE images from ILD patients

The quantitative analyses of CLE images reported includes mainly morphometric measurements made to the fiber structures.

J. Yserbyt [47] [60] has tried measurements such as the alveolar duct diameter, fiber thickness, septal wall thickness, number of alveolar cells per FOV, and the corresponding dimensions and autofluorescence intensity from these cells. He also measured the autofluorescence intensity in the central airways. In one of his studies, a test-retest reliability was performed, using data from 2 consecutive CLE procedures which had an interval of 90 days. The results showed that test-retest reliability is good for number of cells and cells autofluorescence quantification, but only moderate for morphometric analysis of elastic fibers and alveolar ducts.

The same author compared morphometric measurements from the alveolar ducts, as well as the autofluorescence intensity, between a patient with emphysema and a healthy subject [61]. He calculated the cross-sectional area of the alveolar openings at the level of the alveolar ducts (ADA) by measuring the longest diameter (LDAD) and perpendicular short axis (SDAD): $ADA = 0.25\pi \times LDAD \times SDAD$. The study concluded the median cross-sectional area of the alveolar openings at the level of the alveolar ducts is significantly larger in emphysema ($7.2 \times 10^4\ \mu m^2$) as compared with healthy subjects ($5.2 \times 10^4\ \mu m^2$) (p=0.0002). Furthermore, normalized autofluorescence intensity histograms showed a decrease in median autofluorescence intensity in the CLE images from emphysema patients. Also, it was noted a correlation between lung function parameters and the CLE morphometry and autofluorescent measurements.

L. Thiberville [43] also performed morphometric measurements in the alveolar space. According to the published results, the thickness of the elastic fibers in the alveoli is $10 \pm 2.7\mu m$ $(mean \pm SD)$, the diameter of the alveolar mouth is $2.78 \pm 53\mu m$ and microvessels measure $90\ \mu m$ in diameter (mean value).

Likewise, P. Meng [56] performed morphometric measurements: diameter and thickness of alveolar mouth, diameter of the microvessels, number of non-alveolar mouth fibers, and, regarding the auto-fluorescent cells, the number, diameter and autofluorescence intensity of the cells (see Figure 2.12). Their sample included 27 patients and they concluded that the measurements of alveolar mouth dimensions and fiber size were of poor "diagnostic utility", when trying to separate the 2 groups they considered - "chronic fibrosing interstitial pneumonia" (16 patients) and "other ILDs".

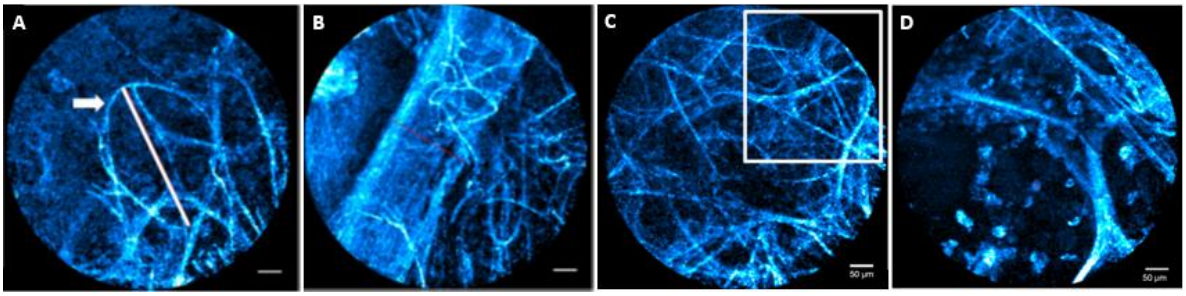

Figure 2.12 – Morphometric measurements in CLE images from ILD patients conducted by P.Meng. A) alveolar mouth diameter (white line) and mouth fiber thickness (arrow), B) microvessel diameter (thin red line), C) counting of non-alveolar-mouth fibers in one quadrant of the FOV, D) counting of cells in one FOV. Scale bars: 50µm. Image from [56] *(adapted)*

From the publications presented above, it can be deducted that qualitative analysis to CLE images in pulmonology are subjective, depending on the observer, and time consuming. In turn, the quantitative methods presented are more objective but are also slow processes since the measurements were made manually using the Cellvizio Viewer software. In fact, there are still few studies using CLE in ILD, and they have small sample sizes which limit the conclusions strength.

2.4.4 Texture Analysis in CLE images

A few research groups, working with Confocal Laser Endomicroscopy imaging in different clinical fields, applied texture analysis to their data.

B. André et al. [62], [63] used shape and texture features to classify automatically CLE images from neo-plastic vs. non-neo-plastic colorectal polyps. They used mosaics to work with larger fields-of-view and retrieved discriminative texture and shape information by applying local operators to those images. The group applied a content-based video retrieval method called bag-of-visual-words (BOW) using an expert-annotated database: the method detects salient image regions from which continuous features are extracted and discretized; then, the features are clustered into a finite number of labels – the visual words –, whose frequencies constitute the image signature. The resultant automated classification of the CLE videos achieved high performance, comparable to the diagnosis established by expert endoscopists who watched and classified the same CLE videos.

C. Désir et al. [64] used shape and texture features to distinguish between healthy and pathological lung tissue. The research group applied and compared the following texture features: SIFT (scale invariant feature transform), Haralick parameters and Local Binary Patterns (LBP). The SIFT extracts robust *keypoints* from the image and describes the neighborhood for each *keypoint* with a histogram of oriented gradients. In turn, the LBP compute the distribution of binary patterns in the circular neighborhood of each pixel. The Haralick parameters are a set of textural features – including, for example energy, entropy and contrast – which are extracted from a matrix (the gray-level co-occurrence matrix, GLCM)

describing the relation between pixel intensities. The classification of patients between healthy and pathological showed a good performance with the best results associated with the Local Binary Patterns, followed by the Haralick parameters extracted from the GLCM and then the SIFT.

A. Rakotomamonjy [65] applied scattering transform – a successive application of a wavelet-based linear filtering – and Local Binary Patterns to discriminate between CLE images from lung cancer patients and healthy subjects. The scattering transform showed a better performance than the LBP, although the combination of both techniques improves the classification. It was concluded that the scattering transform is more suitable for discriminating in coarse scales, while the LBP are more appropriate for discriminating at fine scales.

It is relevant to note that CLE images from different clinical fields (e.g., gastroenterology vs. pulmonology) or diseases (e.g., lung cancer vs. interstitial lung disease) reflect different visual features: while a colorectal polyp is visually distinguished mostly by its shape, the fibrosis in ILD is characterized by the complexity present on the image.

3 MATERIALS AND METHODS

This chapter describes the describes the clinical trial behind this study, the bronchoscopic procedure necessary to record the CLE movies and the subsequent handling of these movies, describing how the numerical features were obtained.

3.1 Patients

From 2016 to 2017, 20 patients with ILD scheduled for diagnostic bronchoscopy with biopsy and/or bronchoalveolar lavage were recruited to participate in the "Confocal Laser Endomicroscopy (CLE) and Optical Coherence Tomography (OCT) for Diagnosing ILD" clinical trial at the Academic Medical Centre (AMC) in Amsterdam [66]. The study took benefit from the fact that the patients had to undergo an endoscopic procedure to make some additional imaging with probe based optical techniques (CLE and OCT). The data used in this project will exclusively be the CLE data.

The eligible participants were above 18 years-old, with both sexes included, a suspected ILD based on HRCT-scan and an indication for bronchoscopy with cryobiopsy. The exclusion criteria were: smoking in the previous 6 months, inability and willingness to provide informed consent and inability to comply with the study protocol.

The 20 patients were diagnosed in multidisciplinary discussions (MDD) at the Academic Medical Center (AMC) in Amsterdam according to the official statement of the American Thoracic Society (ATS), the European Respiratory Society (ERS), the Japanese Respiratory Society (JRS), and the Latin American Thoracic Association (ALAT) [13]. Besides the pulmonologist, radiologist, and pathologist, the multidisciplinary team at the AMC also includes specialists in pulmonary endoscopy, surgery, radiotherapy and oncology. The MDD was done after the patient had underwent the HRCT imaging and the endoscopic procedure with the BAL, CLE (and OCT) and biopsy in some cases.

3.2 Methods

3.2.1 Bronchoscopic Procedure

All patients were subject to a HRCT before they were indicated for bronchoscopy and recruited. The resulting HRCT images were used as reference to know the more affected areas of the lung and the severity of the lung injuries.

The protocol started with the sedation of the patient. Then, the first procedure was the BAL (bronchoalveolar lavage, Figure 3.1). As mentioned in the introduction, the BAL uses a flexible bronchoscope to pump a saline solution into the bronchial tree and then collect it by suction. This was done in only one of the patient lungs.

Figure 3.1 – Bronchoscopic procedure with BAL. Image rights: L. Wijmans, Academic Medical Center, Amsterdam, Netherlands

Afterwards, the CLE protocol took place. The AlveoFlex miniprobe was inserted in the flexible bronchoscope and pushed until the distal airways. The clinician intents to image specially the regions identified as pathological in the HRCT. The laser automatic control was enabled during the procedure, meaning the intensity of the images was adjusted so the structures would have better contrast. This does not allow for a future intensity analysis of the images.

The OCT (optical coherence tomography) protocol was also performed, although its images won't be used for this project. Afterwards, in some cases, the transbronchial lung cryobiopsy (TBLC) is performed to obtain a sample of lung tissue.

The TBLC uses a rigid bronchoscope or tracheoscope. The biopsy was guided with fluoroscopy (the C-arm x-ray system is visible in Figure 3.1). Once the probe was inside the bronchoscope/tracheoscope and reached the desired position, its tip was cooled with carbon dioxide (CO2) for approximately 5–6 seconds until the temperatures reached -75ºC. Then, it was retracted, and the frozen lung tissue stayed attached to the probe's tip. Outside the patient's body, the frozen specimen was submerged into room temperature saline to thaw and then transferred to formalin for fixation. Figure 3.2 shows a cryoprobe before and after being frozen.

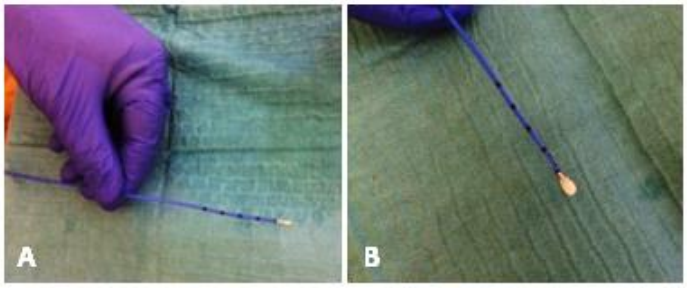

Figure 3.2 – Cryoprobe before (A) and after (B) freezing. Image rights: L. Wijmans, Academic Medical Center, Amsterdam, Netherlands

For some of the patients, before the fixation of the cryobiopsy, the specimen was imaged *ex vivo* with the pCLE probe.

As mentioned before, the patients were diagnosed in multidisciplinary discussions (MDDs) at the Academic Medical Center (AMC) according to the official statement of the ATS, ERS, JRS, and ALAT (American Thoracic Society, European Respiratory Society, Japanese Respiratory Society and Latin American Thoracic Association) [13], with pulmonologists, radiologists and pathologists presenting each patient's case and the available data from the CT-scan, HRCT, BAL and biopsy to be discussed also with specialists in pulmonary endoscopy, surgery, radiotherapy and oncology. The biopsy description made by the pathologist was done according to the normal procedure which includes hematoxylin-eosin (HE) staining.

3.2.2 Handling the CLE Data

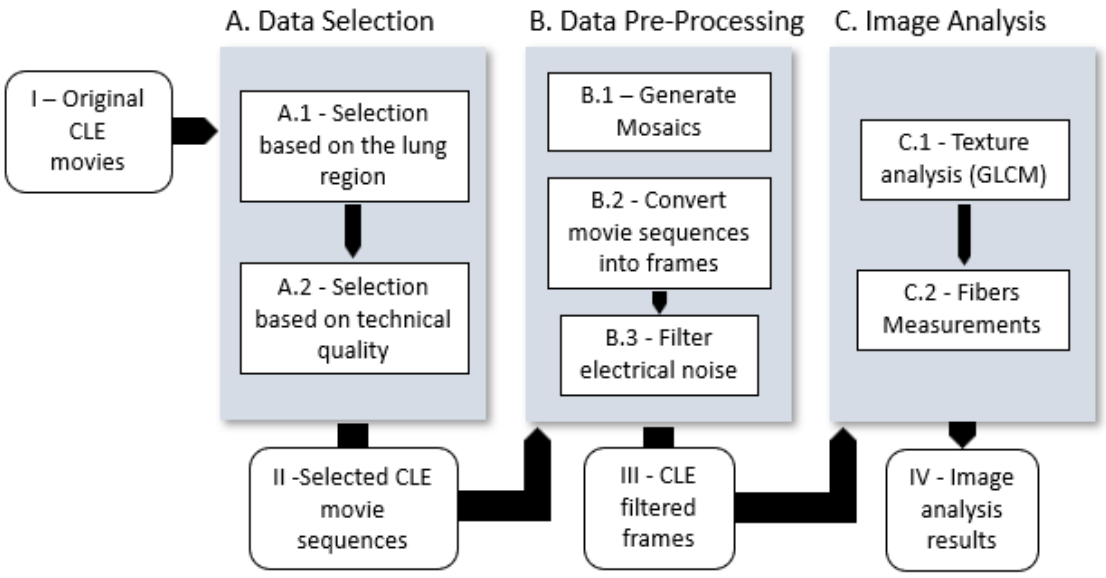

Figure 3.3 – Diagram showing the handling process of the CLE data

The CLE data was processed as shown schematically in Figure 3.3: taking the original CLE movies from the Cellvizio console, they were selected, pre-processed and finally the images were analyzed.

A. Data Selection

Starting with a dataset including more than 160 pCLE movies, a selection was done based on the location of the recording and later based on the technical quality of the movie, as it will be explained bellow.

First, for each patient, were selected the CLE movies which were imaged in the same region from where a biopsy had been taken. This means that, if patient number one (ILD01) had a biopsy taken in the right lower lobe (RLL), only the CLE movies from the RLL were included. This was done so the data from the biopsies could be later used as a reference. The resulting sample of CLE images was relatively big: every patient had approximately 15 minutes of recordings (around 3 videos with 5 minutes each), which lead to a total of 3 hours of CLE movies. If considering the possibility of performing visual scoring over the movies, this amount of time would not reasonable to ask for a scorer to watch and score.

However, because the technical quality of these movies was not always satisfactory, a selection based on technical quality was performed. The movies were assessed by a group of two observers on the following criteria: sharpness of the image, visibility of alveolar structures, condition of the tissue, presence of dirt in the probe or fluid bubbles (from the preceding BAL). It was also registered if there were cells or vessels in the images. To do so, the movies were viewed frame by frame in the Cellvizio Viewer software. For every range of frames which was considered to be technically acceptable, the info on the sequence starting and ending frames was written down, so it could be cropped afterwards. For a range of frames with unacceptable technical quality, besides the start and end frames numbers, it would be noted the reasons for excluding that sequence.

For a sequence to be excluded it had to meet at least one of the following conditions:

a) Lack of alveolar structures. Specifically, less than 75% of the FOV;
b) Stretched tissue due to over pulling the probe, or distorted tissue due to movement artifacts;
c) Unacceptable sharpness of the image;
d) Alveolar septa destruction because of the miniprobe;
e) Presence of fluid bubbles;
f) Excessive dirt on the miniprobe.

B. Data Pre-Processing

The first step when pre-processing the data was to run the algorithm from the Cellvizio Viewer software which searches for the stable video sequences and tries to generate mosaics. To produce a mosaic, the consecutive frames from the stable sequence were "stitched together" resulting in an image with a wider field of view (FOV).

Afterwards, the processing of the CLE dataset (including movies not necessarily stable) took place. The selected CLE movies (originally in the mkt. format) were converted into frames in png format using Matlab (The MathWorks Inc. USA)[67].

Due to electronic interference, some of the CLE video images recorded had a noisy pattern which appears in the images in the form of a grid (see Figure 3.4). To remove that pattern from the image, it is necessary to filter it in the frequency domain using the Fast Fourier Transform (FFT). Because repetitive noise is usually seen in the Fourier spectrum as bright peaks somewhere other than the origin of the frequency domain (dc component), it can be suppressed by erasing those peaks using a notch filter. Indeed, notch filters (or band-stop filters) are like narrow high-pass filters, which notch/cut out frequencies other than the dc component. The frequencies causing the noise (and some of its neighbors) are attenuated and the other frequencies of the Fourier transform are left relatively unchanged.

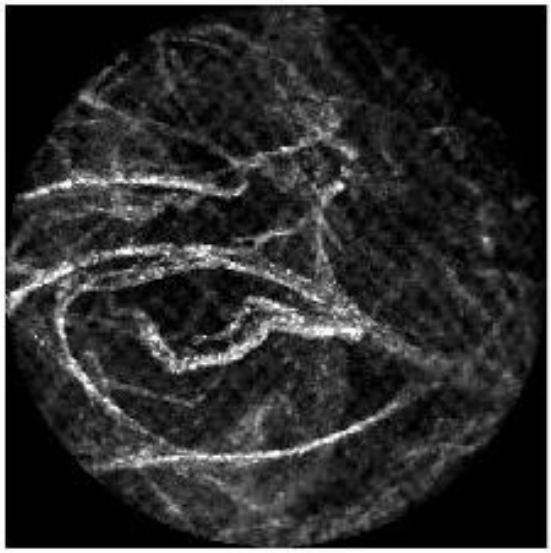

Figure 3.4 – CLE image showing electronic interference in the form of a grid. Scale: the FOV diameter corresponds to 600 µm.

Matlab (The MathWorks Inc. USA [67]) was used to create an algorithm to pre-process the data. The "NotchFilter_CLE" script was written to filter the noise present in the CLE images. If a movie had a noise pattern due to electronic interference, in principle, the pattern would be the same across all the frames. In this sense, only one frame (the "reference frame") would be used to determine the frequencies to be filtered. Then, after designing the filter, it would be applied to all the frames from the referred movie. The script opens the reference frame, converts it to a normalized grayscale format (with intensity values from 0 to 1), and applies zero padding allowing to use a longer FFT with more frequency bins. After the FFT is applied to the image, the noisy frequencies can be identified. This process is done, if possible, using the function written for that purpose "findNoiseFreq.m", which selects the frequencies with the highest FFT values, and ignores the peaks relative to the dc component and the low frequencies associated with the image background. No more than 4 frequencies were selected in each image. Using those values, the notch filter is applied. This filter is a Butterworth highpass filter of order 1, and cutoff frequency 10. If the resulting image still shows the interference pattern, it is necessary to search the intended frequencies manually. In fact, notch filtering is an ad hoc procedure requiring most of the times a human expert to select the frequencies to be removed [42].

Some of the functions called by the "NotchFilter_CLE" Matlab script, namely the ones that calculate the size of the zero padding matrix ("paddedsize.m"), the low pass filter ("lpfilter.m") and the notch filter ("notch.m") were based on the functions present in "Frequency Domain Processing Lab" from the Department of Computer Science, University of Regina [68].

C. Image Analysis

After the image selection and pre-processing, the image analysis took place. This process extracted values from the CLE frames in an attempt to quantify the amount and organization of the elastic fibers present in the alveolar space. Thus, the fibers were segmented, then counted and measured and the image texture was calculated through the Gray-Level Co-Occurrence Matrix (GLCM).

No intensity measurements were used since, as mentioned before, the laser automatic control was enabled, meaning the pixel intensities were automatically adjusted by the Cellvizio system while the movies were being recorded. Therefore, we must not rely on isolated intensity values. Instead relationships between pairs of pixels through GLCM were used.

The measurements were calculated using the FIJI (ImageJ) software: the texture analysis algorithm was based on the plugin "GLCM_Texture_Too" v0.009 from Toby Cornish and Knut Kvaal, while the segmentation of the elastic fibers and subsequent analysis was performed using the "Ridge Detection" v1.4.0 plugin by Thorsten Wagner and Mark Hiner. The latter is based on the model for extracting asymmetrical bar-shaped lines created by Carsten Stege [69]. Further calculations were computed using RStudio [70].

The texture analysis was computed inside a circular ROI circumscribing the FOV area. The following measurements were obtained:

C.1. – Texture Analysis (GLCM):

C.1.1 – Angular Second Moment (Energy)

C.1.2 – Entropy

C.1.3 – Inverse Difference Moment

C.1.4 – Contrast

C.1.5 – Variance

C.1.6 – Correlation

C.2 – From the segmented fibers:

C.2.1 – Number of detected fibers

C.2.2 – Length of the detected fibers

C.2.3 – Width of the detected fibers

C.2.4 – Number of junctions

C.2.5 – Lengths Sum (sum of the lengths of all detected fibers)

It was observed that sometimes the ridge detection algorithm equivocally detected small structures in the image – usually visual artifacts – as fibers. Therefore, it was established that the detected fibers with less than 20 μm would be subsequently excluded from the data. Those were not included in the calculations to obtain the Number of Detected Fibers, nor the Lengths Sum. Also, the junctions involving one of these excluded fibers were omitted from the calculation of the number of junctions.

The final output was the set of statistical variables minimum, maximum, median, 1st and 3rd quartiles extracted from the mentioned variables.

Although the image analysis algorithm looped trough each sequence at a time, the following extraction of the statistical values occurred considering the total "frames set", as if there was no separation between movies or sequences within the same patient. In this sense, the final output is given in terms of patients and not in terms of sequences or movies. This decision is based on the fact that there is no radiologic or histopathological information to match each different movie/sequence, instead the descriptions are associated with the lung region.

For the texture features C.1.1– C.1.6 and variables C.2.1, C.2.4 and C.2.5, one value per frame was computed. If this patient's "frame set" had size F (F frames), then, a list of F values was created for each of those variables. Thus, the statistical variables were extracted from those lists. Conversely, for the variables "Length of the detected fibers" (C.2.2) and "Width of the detected fibers" (C.2.3) there are multiple values per frame (one per detected fiber). In this case, the outputs from each frame were

concatenated into a list (with size equal to the total number of fibers detected in the entire "frame set") and the statistical variables were extracted from that resulting list.

The next sub-chapter will cover texture analysis, specifically, the computation of the GLCM and extraction of its texture features.

3.2.3 Texture Analysis

Texture analysis is a class of mathematical procedures and models which characterize spatial variations within images to extract information. In this project, the Gray-Level Co-Occurrence Matrix (GLCM) introduced by Robert Haralick was used to extract second order statistical textural properties from the CLE images.

The algorithm implemented in ImageJ (FIJI) generated a symmetrical gray-level co-occurrence matrix with spatial displacement between pixels preset as $(d, \theta) = (1,0)$. As mentioned previously, the features extracted were Angular Second Moment (Energy), Inverse Difference Moment, Contrast, Entropy Variance and Correlation.

Next, the steps to compute a GLCM will be described and the equations which define the Haralick texture features will be presented.

3.2.3.1 Compute the GLCM

The Gray-Level Co-Occurrence Matrix represents the relation between pairs of pixels in one image.

If the image I to be analyzed has G gray-levels, then the corresponding GLCM will be a $G \times G$ matrix. The following steps are necessary to create a symmetrical normalized GLCM $P(i,j|\Delta x, \Delta y)$ (or $P(i,j|d,\theta)$), from that image I:

1) Create an $G \times G$ empty matrix, where G is the number of gray-levels from the image I
   Let's call this matrix C.
2) Select the spatial relationship between the reference and neighbor pixel
   When looking at pixel pairs, one of them is called the reference pixel and the other is the neighbor. The chosen relationship between them can be defined in terms of a distance $(\Delta x, \Delta y)$ or a displacement and angle (d, θ).
3) Count the occurrences of pair of pixels with the chosen spatial relationship and fill in the matrix, according to the equation:

$$C(i,j|\Delta x, \Delta y) = \sum_{n=1}^{N-\Delta y} \sum_{m=1}^{M-\Delta x} A \tag{3.1}$$

$$A = \begin{cases} 1, & if\ I(m,n) = i\ and\ I(m+\Delta x, n+\Delta y) = j \\ 0, & elsewhere \end{cases}$$

   An example is given bellow in Figure 3.5
4) Add the matrix C to its transpose to get the symmetrical matrix $S(i,j|d,\theta)$ (see Figure 3.6).
5) Normalize the matrix S to turn it into probabilities.
   Then, the symmetrical normalized GLCM $P(i,j|d,\theta)$ is obtained (see Figure 3.6).

As an example, in Figure 3.5 we have a small image I with dimensions 4×4 which uses 2-bit data, corresponding to $2^2 = 4$ possible gray-levels $G = 4$. Thus, the GLCM will also be a 4×4 matrix. The chosen spatial relationship was $(d, \theta) = (1,0)$ so that the neighbor pixel is displaced by one unit rightward (east) relative to the reference pixel. In this sense, the matrix element $C(i,j|d,\theta)$ is the number of occurrences for having two pixels, separated by (d, θ), one with intensity i and the other with intensity j.

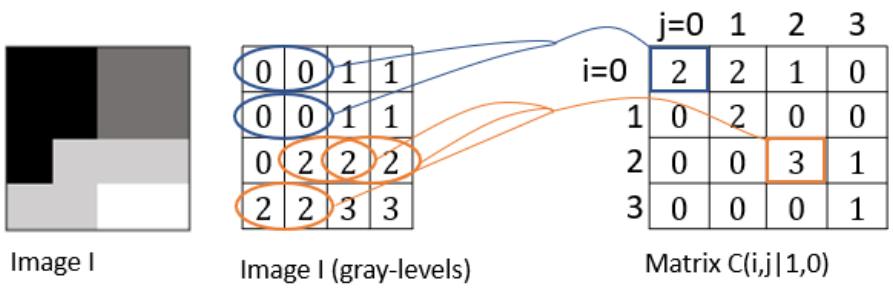

Figure 3.5 – Example of step 3 to get a symmetrical normalized GLCM: Count the occurrences of pair of pixels in the image I with the chosen spatial relationship being $(d, \theta) = (1, 0)$ and fill in the matrix C. On the left is the original 2-bits image I, and the matrix at the center contains the gray-levels associated with each pixel from that image. C is a 4×4 matrix due to the $2^2 = 4$ levels of gray in I. To fill the top left value in the matrix (i=0, j=0) we counted how many times in the image I we have a reference pixel with intensity i=0 whose neighbor has intensity j=0. As it is highlighted in blue, this happens twice, so that C(1,0;1,0)=2.

To have a GLCM symmetrical around the diagonal will help the calculations for the texture features and will diminish the problem with the pixels on the edges of the image, making it possible for every pixel to be a reference and neighbor pixel. In order to make matrix C symmetrical (obtaining matrix S), each pixel pair is counted twice: once "forwards" and once "backwards" (interchanging the reference and neighbor pixels for the second count). Alternatively, the symmetrical matrix S can also be calculated by adding the "original" matrix to its transpose (see Figure 3.6). Besides the symmetry around the diagonal, the GLCM should be expressed as a probability instead of absolute counts of occurrences. Therefore, each value of the matrix $S(i,j|1,0)$ should be divided by the sum of all its entries, which is 24 in this example. As a result, the elements of the GLCM may now be considered probabilities of finding 2 pixels with intensities (i,j) with a particular spatial relationship in the image I. The resulting symmetrical normalized GLCM $P(i,j|1,0)$ is shown in Figure 3.6.

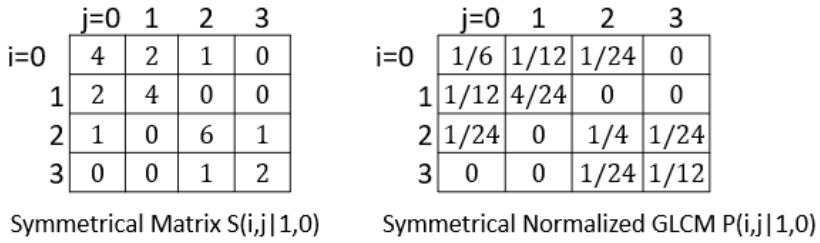

Figure 3.6 – Example of the symmetrical (left) and symmetrical normalized (right) gray-level co-occurrence matrices

After the computing the GLCM, the texture features can be extracted from it.

3.2.3.2 Texture Features to be extracted from the GLCM

From the set of 14 features proposed by Haralick to be extracted from the GLCM, the following 6 were applied in this project: Angular Second Moment (Energy), Inverse Difference Moment, Contrast, Entropy, Variance and Correlation [71], [72].

The variables presented bellow in Table 3.1 with corresponding equations (3.2) to (3.9) will be necessary to calculate the texture features:

Table 3.1 – Variables used in the gray-level co-occurrence matrices

G is the number of gray levels used
μ is the mean value of P
μ_x and μ_y are the means of P_x and P_y
σ_x and σ_y are the standard deviations of P_x and P_y
$P_x(i)$ is the ith entry in the marginal-probability matrix obtained by summing the rows of $P(i,j)$

$$P_x(i) = \sum_{j=0}^{G-1} P(i,j) \tag{3.2}$$

$$P_y(j) = \sum_{i=0}^{G-1} P(i,j) \tag{3.3}$$

$$\mu_x = \sum_{i=0}^{G-1} i \sum_{j=0}^{G-1} P(i,j) = \sum_{i=0}^{G-1} iP_x(i) \tag{3.4}$$

$$\mu_y = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} jP(i,j) = \sum_{i=0}^{G-1} jP_y(j) \tag{3.5}$$

$$\sigma_x^2 = \sum_{i=0}^{G-1} (i-\mu_x)^2 \sum_{j=0}^{G-1} P(i,j) = \sum_{i=0}^{G-1} \big(P_x(i) - \mu_x(i)\big)^2 \tag{3.6}$$

$$\sigma_y^2 = \sum_{j=0}^{G-1} (i-\mu_y)^2 \sum_{i=0}^{G-1} P(i,j) = \sum_{j=0}^{G-1} (P_y(i) - \mu_y(i))^2 \tag{3.7}$$

and

$$P_{x+y}(k) = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i,j) \tag{3.8}$$

$i + j = k$, for $k = 0,1, \dots, 2(G-1)$.

$$P_{x-y}(k) = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i,j) \tag{3.9}$$

$|i + j| = k$, for $k = 0,1, \dots, (G-1)$.

Now, the 6 chosen Haralick texture features will be described.

The Angular Second Moment (ASM) – sometimes referred as Energy/Homogeneity/Uniformity by other authors – measures the homogeneity of the image and it is calculated as the sum of the squares of the GLCM entries:

$$ASM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\{P(i,j)\}^2 \tag{3.10}$$

In a homogeneous image, there are very few dominant gray-level transitions, so the GLCM will have fewer entries of large magnitude and the ASM will be higher. On the other hand, for an inhomogeneous image, the GLCM will have large number of small entries, thus the ASM feature will be smaller.

Inverse Difference Moment (IDM) - sometimes called Local Homogeneity by other authors - is calculated by:

$$IDM = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{1}{1+(i-j)^2}P(i,j) \tag{3.11}$$

This feature is influenced by the homogeneity of the image since the weighting factor $1+(i-j)^{2^{-1}}$ will have small contributions from inhomogeneous areas where $i \neq j$, which means low IDM for inhomogeneous images. IDM is approximately inversely correlated with contrast.

Contrast measures the amount of local variations present in an image. Entries away from the diagonal of the GLCM will have higher contributions to the contrast feature. A high contrast may indicate the presence of edges, noise or wrinkled textures:

$$CONTRAST = \sum_{s=0}^{G-1} s^2 \left\{\sum_{i=1}^{G}\sum_{j=1}^{G} P(i,j)\right\} \qquad |i-j| = s \tag{3.12}$$

Entropy measures the degree of disorder in the image. It is defined as:

$$ENTROPY = -\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P(i,j) \times \log(P(i,j)) \tag{3.13}$$

More complex images, or images with a larger number of gray levels, will have higher values of entropy.

Variance is a measure of heterogeneity. It measures the dispersion (with regard to the mean) of the gray level distribution, by attributing relatively high weights on the elements that differ from the average value of $P(i,j)$. Variance increases when the gray values differ from their means. While a high variance suggests a high contrast, the opposite does not apply.

$$VARIANCE = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}(i-\mu)^2 P(i,j) \tag{3.14}$$

The Correlation feature measures gray-level linear-dependencies in the image:

$$CORRELATION = \sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{\{i \times j\} \times P(i,j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y} \tag{3.15}$$

When comparing different images, it is possible for a human to have some intuitive expectations regarding some of the textural features, such as expecting higher entropy for more complex images or high correlation values when linear dependencies are detectable. However, this is not possible in every textural feature, which also justifies the need for the computational analyses.

4 RESULTS

In this section the patient group will be presented, as well as the results extracted from the CLE images and the corresponding analysis.

4.1 Patient group

Twenty patients were enrolled in the clinical trial from 2016 to 2017 and identified as ILD01 to ILD20.

After the bronchoscopic procedure was concluded, the patients' cases were discussed in the multidisciplinary discussions (MDD). There were 2 patients classified with unknown diagnosis and the remaining 18 patients comprised 10 different types of ILD, which are listed below (according to the scheme shown in the Figure 2.1 from the Introduction):

- ILD of known cause: Silicosis, Chronic EEA (extrinsic allergic alveolitis / hypersensitivity pneumonitis);
- Idiopathic interstitial pneumonias (IIPs): idiopathic pulmonary fibrosis (IPF), desquamative interstitial pneumonia (DIP), fibrotic non-specific interstitial pneumonia (fNSIP), cellular non-specific interstitial pneumonia (cellular NSIP), lymphocytic interstitial pneumonitis (LIP);
- Granulomatous ILD: Sarcoidosis;
- Other forms of ILD: Langerhans' cell histiocytosis (LCH), Pulmonary Alveolar Proteinosis (PAP).

4.1.1 Patient data

Table 4.1 shows the number of patients for each diagnosis, the available data and the patients which had to be excluded from this project. There is HRCT, BAL and CLE data from all the 20 subjects, though biopsies were only taken from 11 patients. Eight patients had to be excluded from this study because of the following reasons: there was no alveolar tissue visible on the CLE images (1 patient), the final diagnosis was unknown (2 patients) and absence of biopsies, therefore no pathological reference (5 patients).

Table 4.1 – Available data from the 20 participants with corresponding diagnosis. The excluded patients are in orange and have the reason for exclusion mentioned. In the Biopsy column is expressed if the procedure was a TLCB (Transbronchial Lung Cryobiopsy) or a surgery VATS (Video-Assisted Thoracoscopic Surgery)

Diagnosis	No. patients	HRCT	BAL	CLE	Biopsy	Reason for Exclusion
DIP	1	x	x	x	TLCB	
Fibrotic NSIP	5	x	x	x	TLCB	
Cellular NSIP	1	x	x	x	TLCB	
LIP	1	x	x	x	VATS	
Chronic EEA	2	x	x	x	TLCB /VATS	
Silicosis	1	x	x	x	TLCB	
PAP	1	x	x	x	TLCB	No alveolar tissue on CLE
Sarcoidosis	3+1	x	x	x	No	No histopathology (3)
LHC	1	x	x	x	No	No histopathology
IPF	1	x	x	x	No	No histopathology
Unknown	2	x	x	x	TLCB	Unknown final diagnosis
TOTAL	20 (12 + 8)					

For the patients which had previously suspected Sarcoidosis by the visible enlarged lymph nodes on the HRCT imaging, their lymph nodes were analyzed with needle-based CLE (nCLE) and the Sarcoidosis diagnosis was confirmed. The process consists in puncturing the lymph node with the nCLE, retract the needle slowly while imaging, and obtain cytologic aspiration which is later analyzed. Because the disease was confirmed by that method, there was no need to subject these patients to a biopsy since it is an invasive procedure. However, one of the four patients which has Sarcoidosis as their final diagnosis was included since he underwent a biopsy to find out his diagnosis.

There was one patient (ILD02) which was initially included in the study, despite not having histopathological data. Since its data was also analyzed and there was a HRCT description, this patient was included in the analysis where CLE was compared with the HRCT outcome.

In conclusion, from the 20 subjects recruited for the clinical trial, only 12 were included in this study. The available data is the HRCT images, CLE videos (*in vivo* and some *ex vivo*), the cytological description from the BAL, and, for those who had a biopsy taken, the corresponding pathological description.

4.2 CLE Data

4.2.1 Visual Analysis of CLE movies

Before making any quantitative analysis to the images, the original CLE movies were watched several times to get a better understanding of the data. The Appendix I shows two examples of CLE movies.

Figure 4.1 shows some CLE images from 5 patients whose diagnosis is fibrotic nonspecific interstitial pneumonia (fNSIP), a subtype of NSIP – NSIP is the second most common ILD, preceded by IPF. The histopathological description from these patients mentioned thickening of the alveolar septa and some patients have additional fibers referred. The HRCT showed the presence of ground-glass opacities, except for patient ILD05. In the figure is possible to observe the heterogeneity within the same patient and between patients with the same disease.

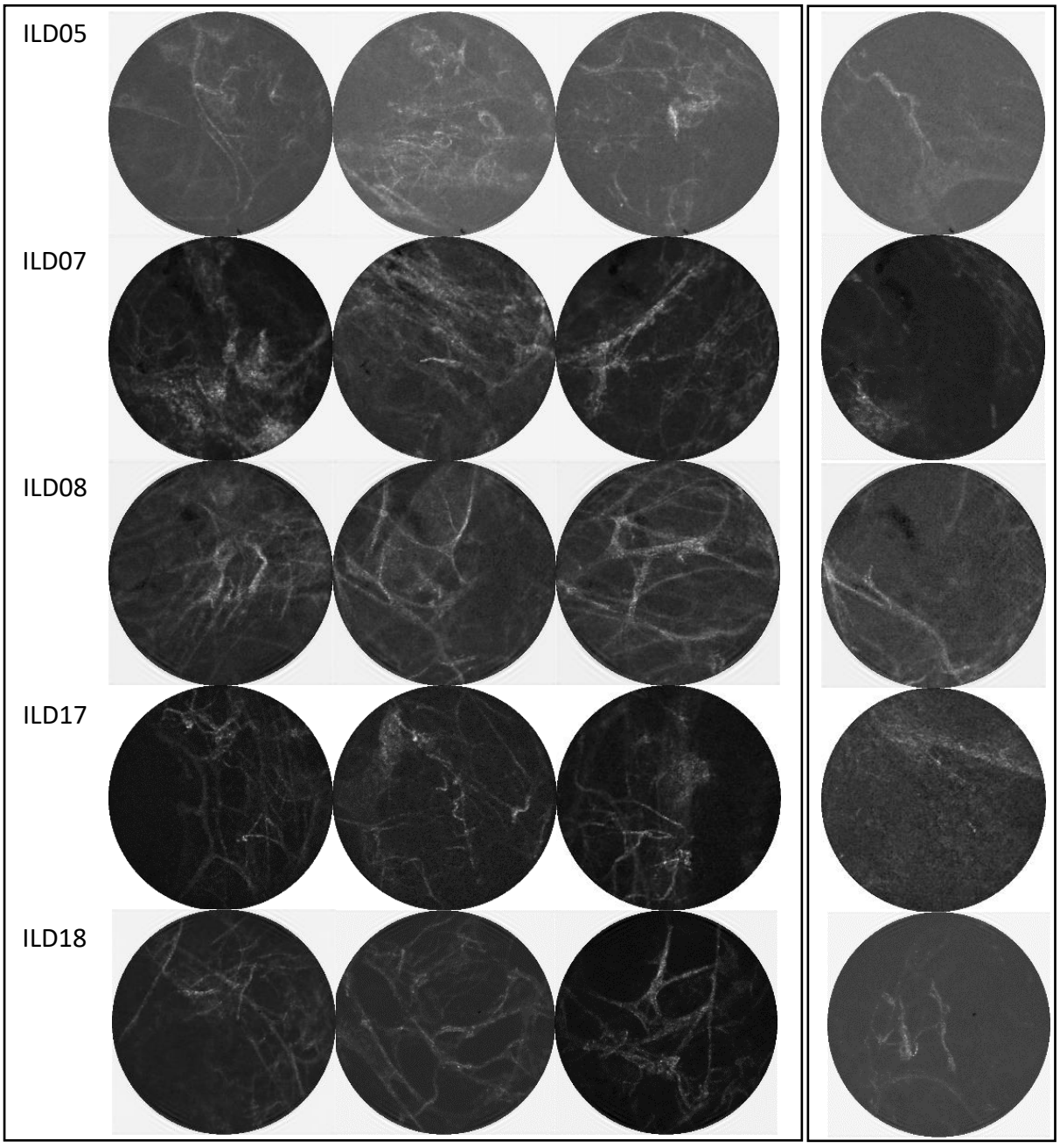

Figure 4.1 – CLE images from 5 ILD patients diagnosed with fibrotic nonspecific interstitial pneumonia (fNSIP). Scale: the FOV diameter corresponds to 600 µm.

While the first 3 columns of images in Figure 4.1 show alveolar spaces with multiple elastin fibers, the last column (separated by the box) shows less fibers or poor contrast, which is a common situation revealed in the data set. Regarding the images on the left, the heterogenic patterns include: normal elastin fibers with healthy architecture, disorganized fibers, thickened fibers and extra fibers comparing to healthy alveoli. Because the CLE laser automatic control adjusts the image intensity during the procedure, the contrast between images varies, even within the same patient.

The images in Figure 4.2 belong to patient ILD11 diagnosed with the other subtype of NSIP: cellular NSIP (cNSIP), which is less common than the fibrotic subtype but has a better prognosis. The HRCT from this patient had diffuse ground-glass opacity and the histopathological description mentions the presence of cellular infiltrates and parts with normal lung tissue. In the CLE images shown below, the white spots are the inflammatory cells characteristic from this disease. Nevertheless, the fibers architecture seen on these CLE images does not appear normal, which may result from the disease state, or the CLE probe movements.

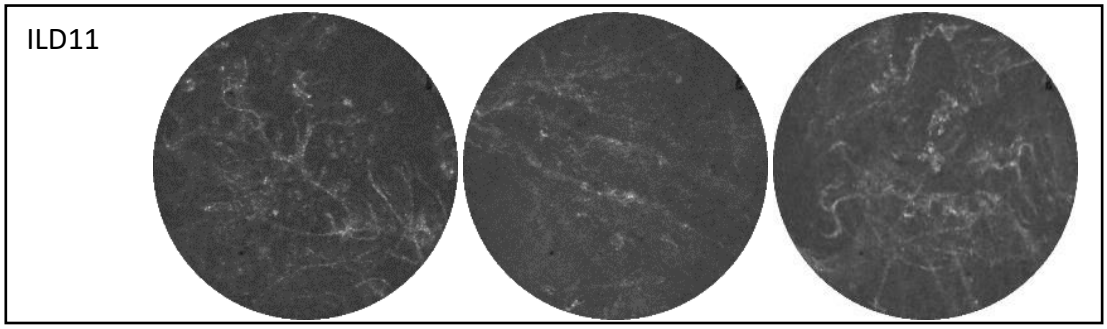

Figure 4.2 – CLE images from patient ILD11 diagnosed with cellular nonspecific interstitial pneumonia (cNSIP). Scale: the FOV diameter corresponds to 600 µm.

Using the Cellvizio Viewer software, we were able to generate mosaics from some video sequences considered stable. These images are very useful for visual analysis because they represent a small movie sequence in one static image. For a mosaic to be used in quantitative analysis, the area of its FOV should be calculated, so measurements like number of elastic fibers can be normalized. To apply GLCM analysis to the mosaics is hard due to the irregular shape of the FOV.

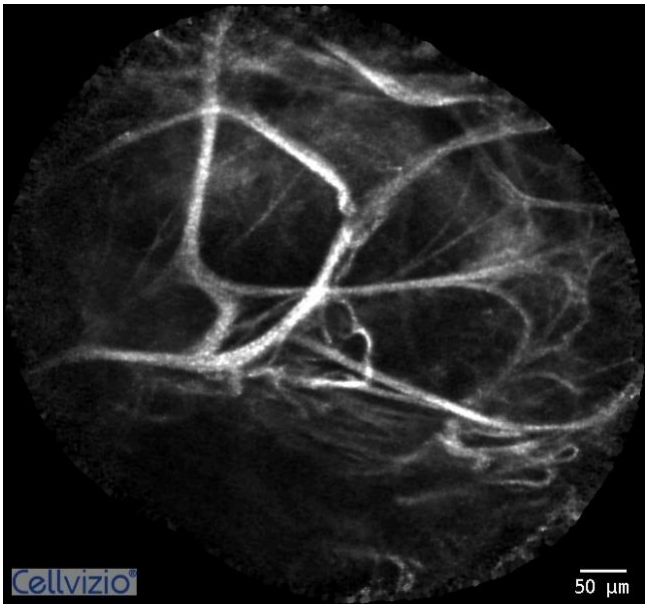

Figure 4.3 – CLE image from patient ILD16 diagnosed with LIP (Lymphocytic Interstitial Pneumonitis). Mosaic obtained using the Cellvizio Viewer software. Scale bar: 50 µm.

Figure 4.3 is a mosaic representing what was imaged in a segment of the alveoloscopy in patient ILD16, diagnosed with LIP (Lymphocytic Interstitial Pneumonitis), normally characterized by diffuse infiltration of the interstitium and alveolar spaces by lymphocytes and plasma cells. The image shows the elastin fibers of the alveolar septa and surrounding alveolar airspaces with some white opacities. Indeed, this patient's HRCT shows ground-glass opacities and the histopathological description mentions the presence of alveolar filling with neglectable fibrosis. However, this description also mentions the destruction of the alveolar network by cellular infiltrates with lymphocytes and giant cells, which is not visible in the CLE sequences.

The following figures (Figure 4.4-Figure 4.5) contain three mosaics from patient ILD09 diagnosed with Chronic EAA (extrinsic allergic alveolitis). The CLE movie sequences that gave rise to these mosaics were imaged *ex vivo*: after the biopsy was collected, the specimen was submerged into room temperature saline to thaw and then imaged with the pCLE probe. This patient HRCT presented honeycombing and the histopathological data showed parts with thin alveolar septa and others with dense fibrosis. The CLE images show some thick and extra elastin fibers but with the elastin architecture preserved. These images have a very good contrast and the alveolar spaces look very clean probably because the biopsy was previously submerged in the saline solution. Also, there is no reference to presence of alveolar filling in the histopathological description. It is possible to obtain bigger mosaics when imaging *ex vivo* because the clinician has more control over the probe.

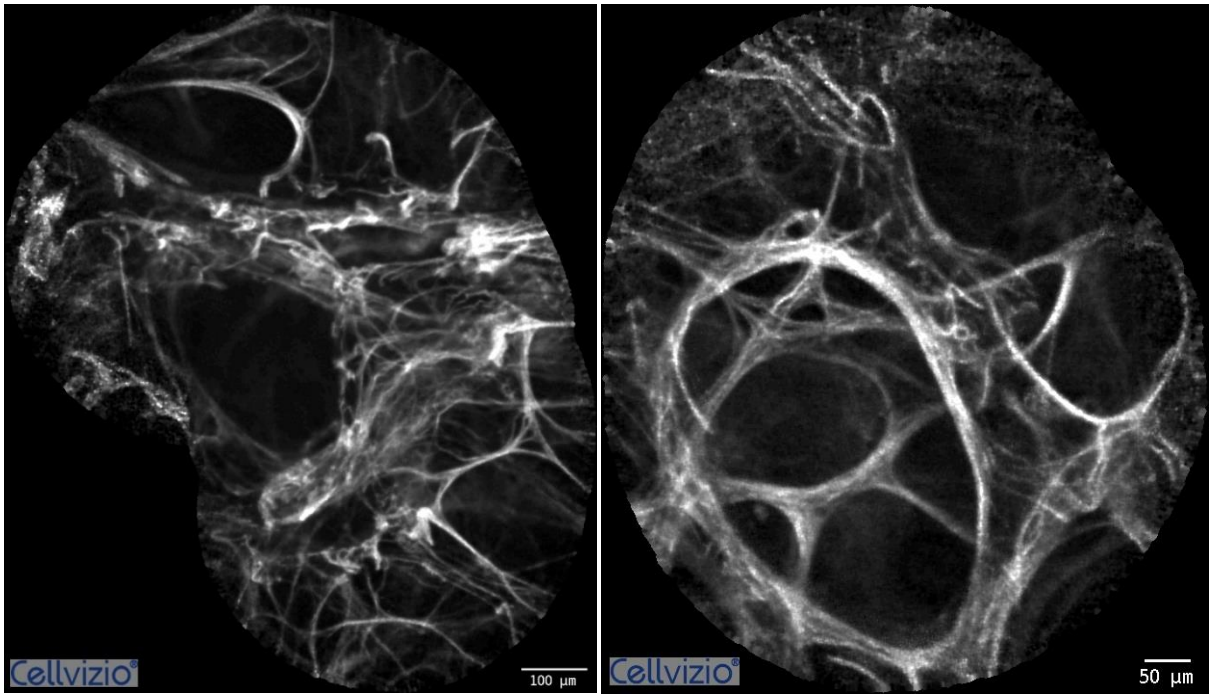

Figure 4.4 – Alveolar tissue ex vivo from ILD09. Mosaics obtained using the Cellvizio Viewer software. Scale bars: 100 µm (left) and 50 µm (right).

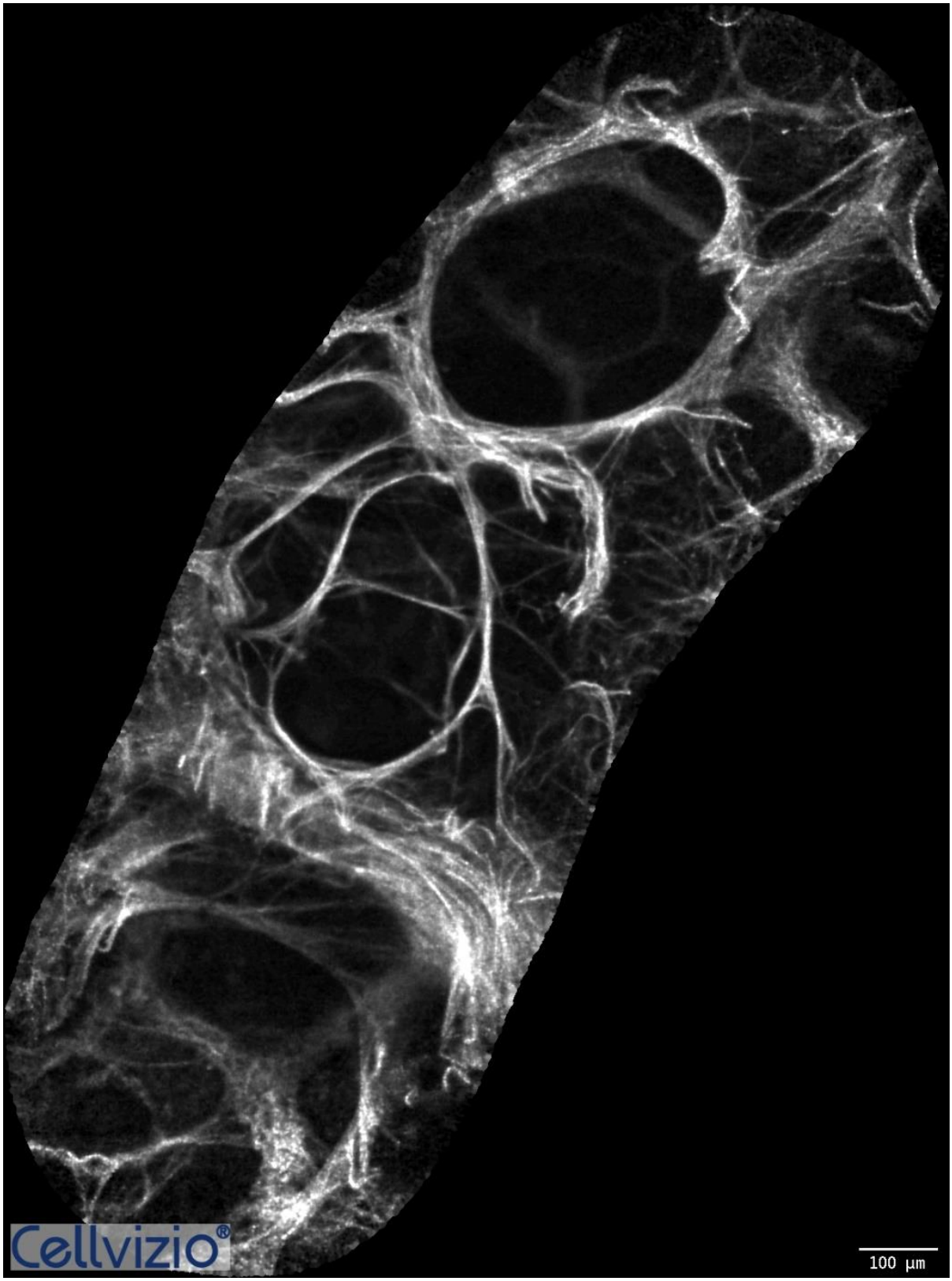

Figure 4.5 – Alveolar tissue ex vivo from ILD09. Mosaic obtained using the Cellvizio Viewer software. Scale bar: 100 µm.

4.2.2 Handling CLE data: one patient's example

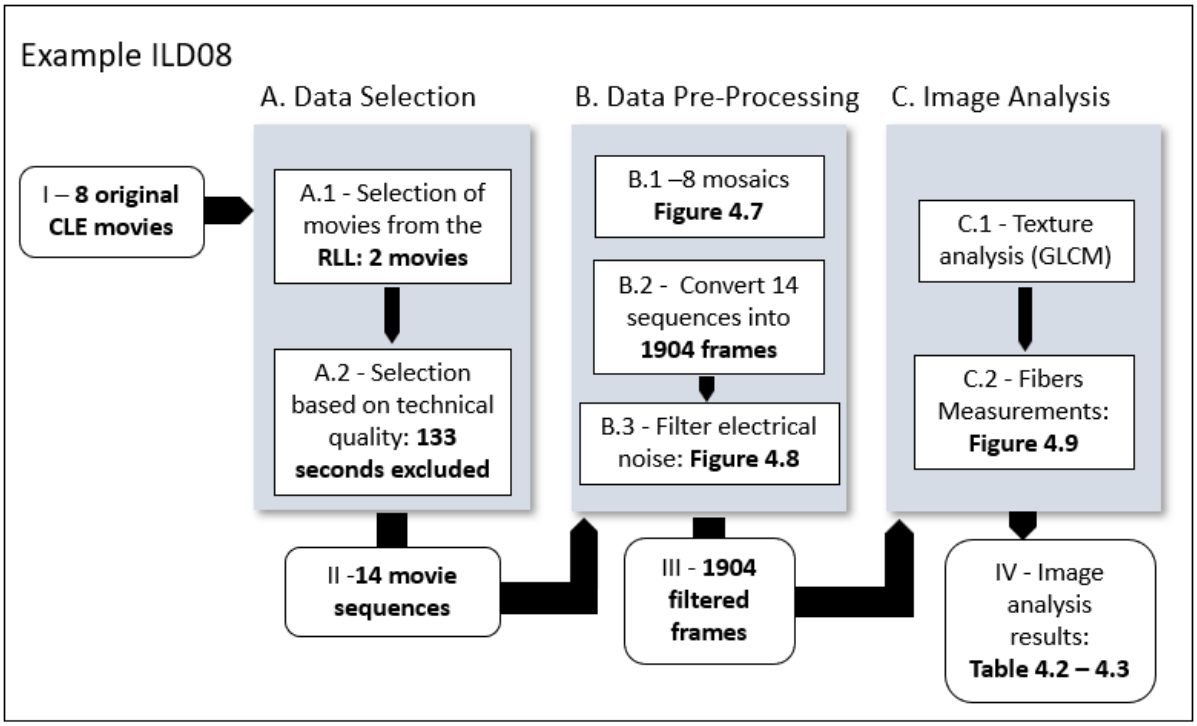

Figure 4.6 – Diagram showing the handling process of the CLE data applied to patient ILD08 (used as an example)

After visualizing the CLE movies, the data was handled for the quantitative image analysis as described in the methods. Using the patient ILD08 as an example, Figure 4.6 represents its data handling process, described below in more detail:

I – 7 CLE movies were recorded during the bronchoscopic procedure.

A. Data Selection

A.1 – This patient's biopsies were collected from the right lower lobe (RLL). Only 2 CLE movies had this location. These 2 movies had originally 3103 frames in total.

A.2 – The selection based on technical quality excluded 1199 frames (approximately 133 seconds of movies) mainly because of lack of visible alveolar structures, resulting in 14 included movie sequences.

II – 14 movie sequences were included.

B. Data Pre-Processing

B.1 – 13 mosaics were generated from the stable movie sequences, but only 8 of them had potential clinical interest. Figure 4.7 shows 2 examples of the mosaics obtained (note that, at this point, the interference pattern had not been filtered from the images yet).

B.2 – The movie sequences were converted into 1904 single frames in png format (see Figure 4.9 A), so the frame set had size 1904.

B.3 – Because the images from this patient had noisy pattern in the form of a grid, 4 frequency peaks associated with that noise were identified and filtered in the frequency domain as seen in Figure 4.8.

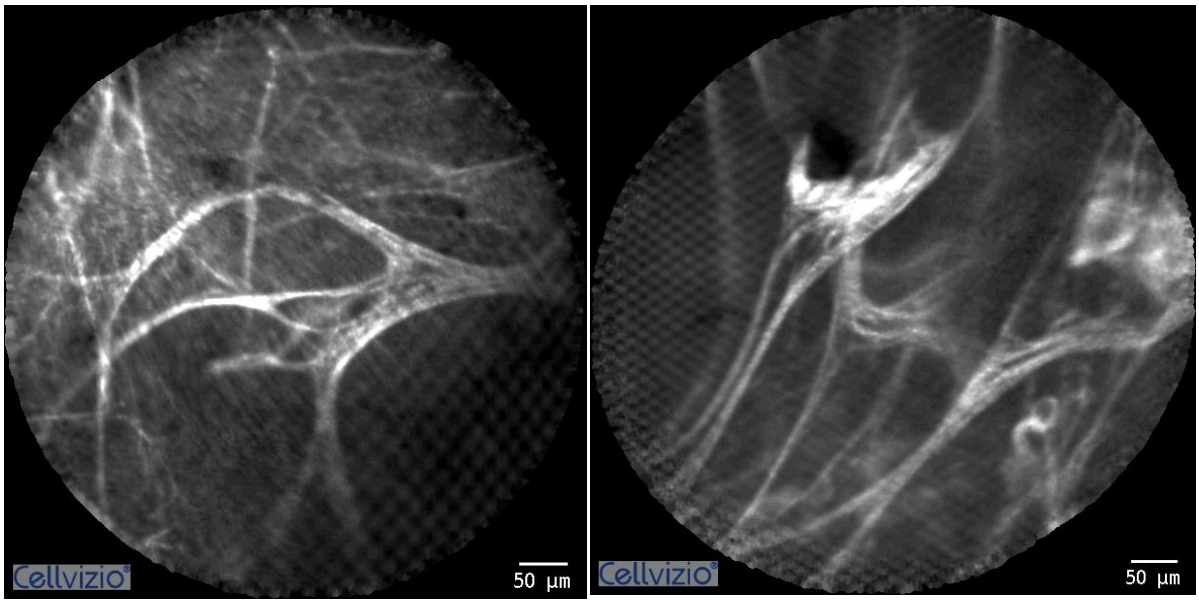

Figure 4.7 – Example of step B.1 "Generate Mosaics" from the Data Pre-Processing: 2 mosaics obtained for patient ILD08 using the Cellvizio Viewer software. Note that, at this point, the interference pattern had not been filtered from the images yet. Scale bars: 50 µm.

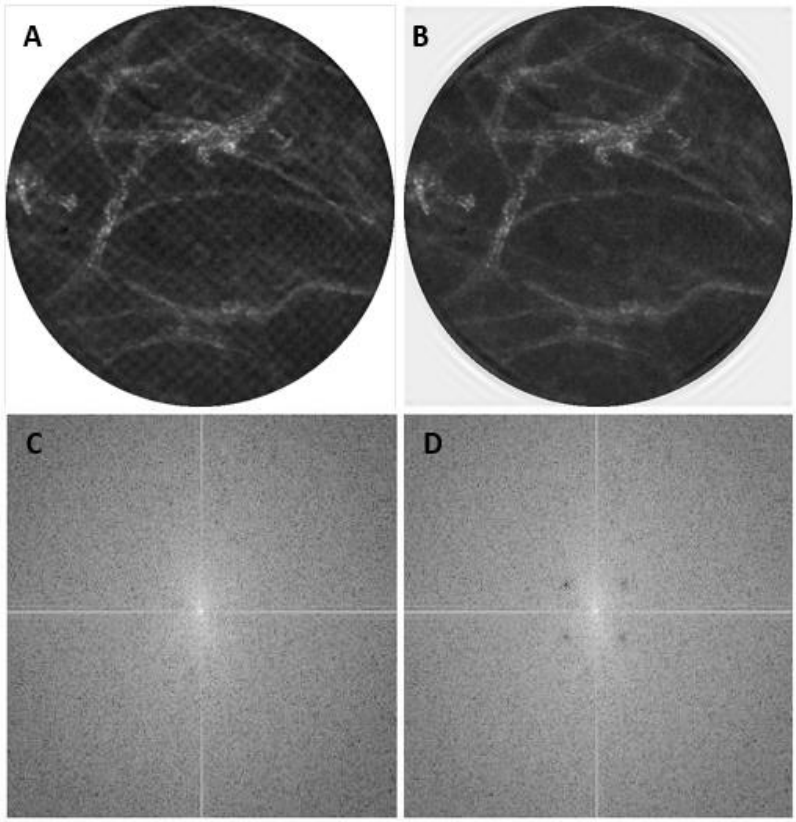

Figure 4.8 – Example of step B.3 "Filter electrical noise" from the Data Pre-Processing. A) The original CLE frame with a noisy pattern visible as a grid and C) corresponding FFTshift plot. B) filtered CLE image and D) corresponding FFTshift plot with de filtered frequencies visible as 4 darker spots. Scale for the CLE images: the FOV diameter corresponds to 600 µm.

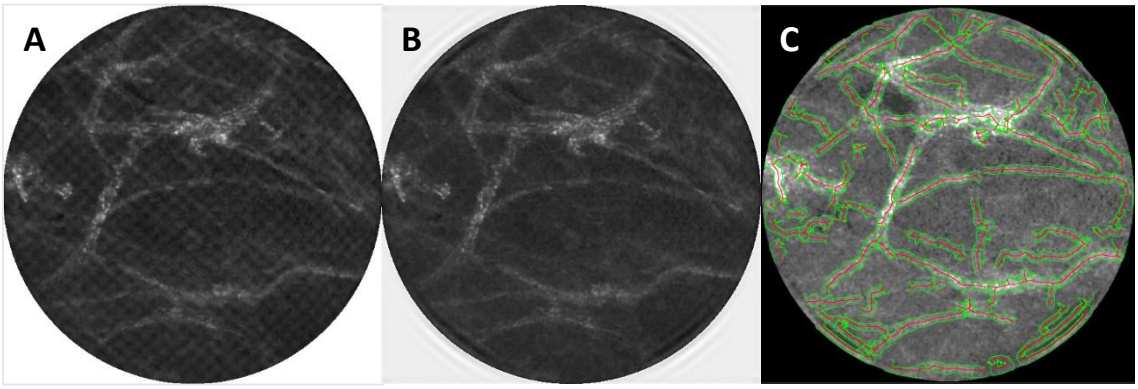

Figure 4.9 – CLE images from patient ILD08 diagnosed with fNSIP. A- original frame extracted from a movie sequence. Notice the noisy pattern with the "grid shape"; B- filtered frame without the noisy pattern; C- Ridge detection algorithm: the red lines represent the detected ridges/fibers and the green lines represent their estimated width. Scale: the FOV diameter corresponds to 600 µm.

II – The 1904 filtered frames were ready to undergo the image analysis process.

C. Image Analysis

C.1 – The GLCM algorithm was computed over every filtered frame. The resulting output can be seen in Table 4.2:

Table 4.2 – Example of the texture measurements obtained for patient ILD08

ASM					**ENTROPY**				
min	1st quart	Median	3rd quart	max	min	1st quart	Median	3rd quart	max
0.001	0.003	0.004	0.005	0.015	4.857	5.852	6.122	6.360	7.490
INVERSE DIFFERENCE MOMENT					**CONTRAST**				
min	1st quart	Median	3rd quart	max	min	1st quart	Median	3rd quart	max
0.225	0.354	0.387	0.421	0.584	2.905	8.135	10.653	13.308	42.428
VARIANCE					**CORRELATION**				
min	1st quart	Median	3rd quart	max	min	1st quart	Median	3rd quart	max
38.190	112.589	149.796	190.976	499.257	0.002	0.005	0.006	0.009	0.025

C.2 – The ridge detection algorithm segmented the fibers as seen in Figure 4.9 C and detected junctions between fibers. It measured the length and width (according to the green lines traced in Figure 4.9 C) for each detected fiber. The detected fibers with less than 20 µm and corresponding junctions were excluded. The Lengths Sum, Number of Detected Fibers and the Number of Junctions were then calculated. The resulting output is in Table 4.3:

Table 4.3 – Example of the fiber measurements obtained for patient ILD08

NO. OF DETECTED FIBERS					**DETECTED FIBERS LENGTH**				
min	1st quart	Median	3rd quart	max	min	1st quart	Median	3rd quart	max
23.00	81.00	98.00	115.00	168.00	20.00	28.80	41.50	65.00	455.20
LENGTHS SUM					**DETECTED FIBERS WIDTH**				
min	1st quart	Median	3rd quart	max	min	1st quart	Median	3rd quart	max
1800.30	4745.60	5277.30	5742.70	6838.10	2.31	10.98	12.16	13.48	24.76
NO. OF JUNCTIONS									
min	1st quart	Median	3rd quart	max					
2.00	33.00	74.00	167.00	350.00					

4.2.3 Results from the entire data set

The selection of video sequences, from the 12 included patients, based on technical quality resulted in 97 sequences with varying lengths making a total of 5854 included frames, which represents almost 11 minutes of recordings in the CLE frame rate (9 fps). Each patient had an average of 8 sequences included, with the minimum number of sequences per patient being 1 and the maximum 14. The average number of frames in one sequence was 60 (almost 7 seconds of recording).

The most common reason to exclude parts of the movies was the lack of alveolar structures (154 excluded sequences), followed by unacceptable sharpness of the image (49 excluded sequences) and stretched tissue due to over pulling the probe, or distorted tissue due to movement artifacts (30 excluded sequences), corresponding to conditions a), c) and b) from section 3.1.2.1, respectively. Only a few sequences were excluded due to alveolar septa destruction caused by the miniprobe, excessive dirt on the miniprobe or presence of fluid bubbles.

The ridge detection algorithm detected most of the elastin fibers recognized by the human eye. Sometimes, some fibers were not detected, probably due to poor contrast or image quality. Because some of the identified ridges didn't correspond to elastin fibers (were visual artifacts, mostly), it was established in the methods that detected ridges with length inferior to $20\ \mu m$ would be excluded from the data. Conversely, sometimes one fiber was misidentified as multiple smaller fibers. The measurement Lengths Sum can be used as a way to get around this problem.

Regarding the texture analysis, the algorithm was run over the dataset, the Gray-Level Co-Occurrence Matrix was computed, and the texture features were calculated. The results presented here are associated with a single spatial displacement $(d, \theta) = (1,0)$. Other distance and angles were also tested but were not included in these results. Although, this was not done due to time constraints, it would be interesting to present the average of a set of co-occurrence matrices for more possible distances and angles, which could give a more complete and isotropic GLCM.

Given the sample size in this study, it was not expected from this data set to allow patient stratification. However, we still chose to present the plots splitting the patients according to the data from the radiologic/histopathological descriptions in the attempt to identify the more promising variables to be used in classification in the future. Thus, multiple graphs were plotted to represent the results from both the texture analysis and the fibers measurements, visually separating the patient groups. It was considered the presence/absence of ground-glass opacities (GGO) described in the HRCT and the presence/absence or fibrosis described in the histopathology, giving rise to the groups "GGO-Y vs. GGO-N" and "Fibrosis-Y vs. Fibrosis-N", with "Y"/"N" meaning yes/no. In a very simplistic definition, the ground-glass opacities present in ILD patients' HRCT scans can be associated with active inflammation or fine fibrosis, and the fibrosis in the histopathology means the deposition of extra and usually disorganized fibers.

Patient ILD02 doesn't have a histopathological description, therefore being excluded from both Fibrosis-Y vs. Fibrosis-N groups.

The boxplots in Figures Figure 4.11 and Figure 4.10, show heterogeneity in the number of fibers detected by the algorithm for each patient. The median values for this measurement lie between 27 and 119 fibers. However, there is no evident difference between the predefined groups of patients in terms of number of fibers. Patient ILD09 whose biopsies were classified has being densely fibrotic, has the greatest interquartile range and the highest maximum value (187 fibers) of Number of Detected Fibers as seen in the boxplots from Figure 4.10. However, ILD07 which is also a densely fibrotic patient does not show a larger interquartile range when compared to other patients.

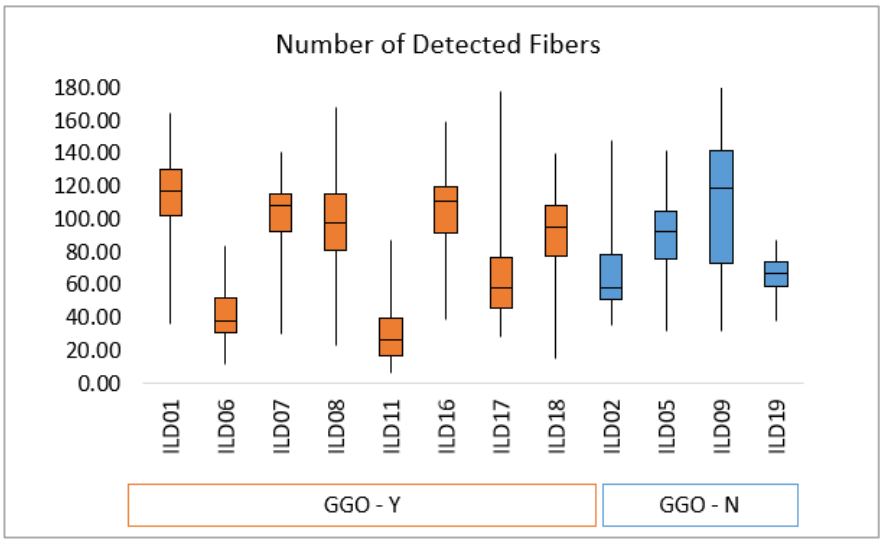

Figure 4.11 – Boxplots representing the Number of detected fibers for each patient. Patients with ground-glass opacities (GGO) described in their HRCT appear in orange (left), and patients without GGO appear in blue (right).

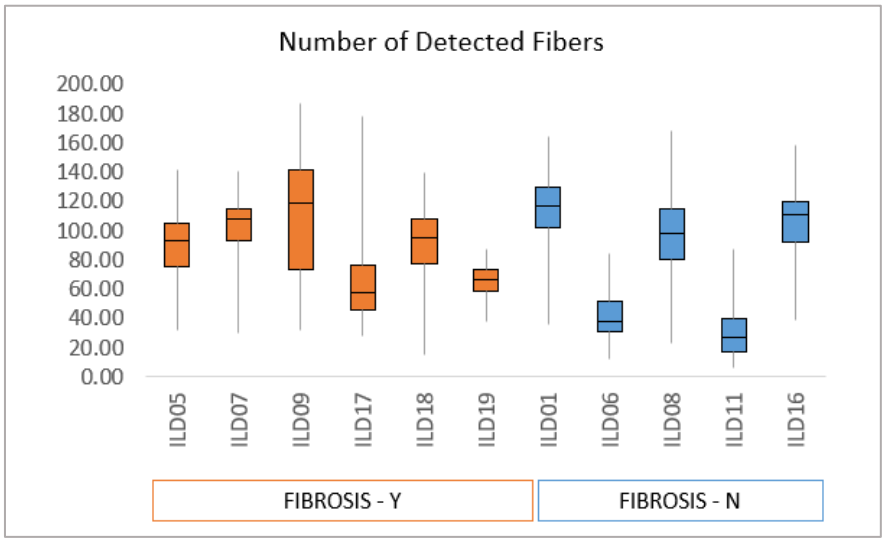

Figure 4.10 – Boxplots representing the Number of detected fibers for each patient. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right).

The Number of Detected Junctions and Number of Detected Fibers showed a positive correlation, as expected. When selecting the fibrotic patients, and visually separating the mild fibrosis from dense fibrosis, a raise both in the Number of Detected Fibers and Junctions can be observed for the densely fibrotic patients (see graph in Figure 4.12).

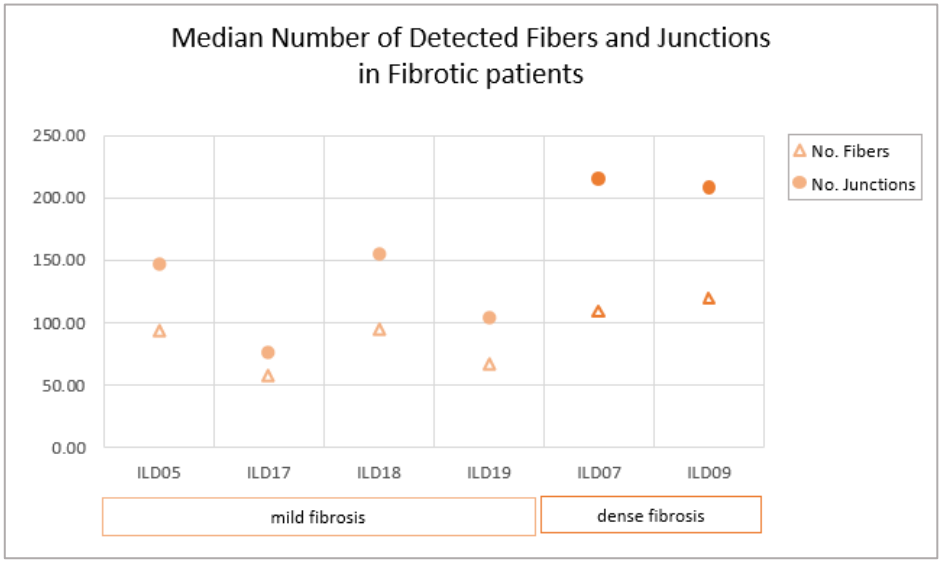

Figure 4.12 – Median values for the Number of detected Fibers (triangles) and Junctions (circles) in the subgroup of patients described as having fibrosis in their biopsies. The mild fibrosis is represented in light orange (left) and dense fibrosis appears darker (right)

In terms of the Length of the Detected Fibers, the median values vary from 34.70 μm and 49.95 μm, and the boxplots in Figure 4.13 show large interquartile ranges. The Lengths Sum measurements, which in some way combine the variables Number and Length of detected fibers, show considerable heterogeneity in the statistical values represented in Figure 4.14.

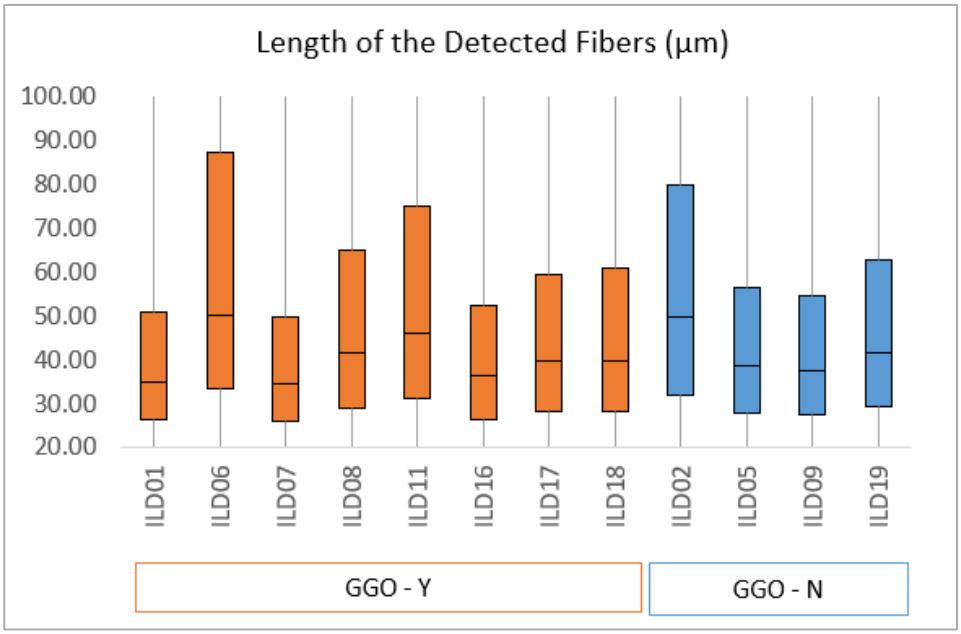

Figure 4.13 – Boxplots representing the length of detected fibers for each patient in µm. Patients with ground-glass opacities (GGO) described in their HRCT appear in orange (left), and patients without GGO appear in blue (right).

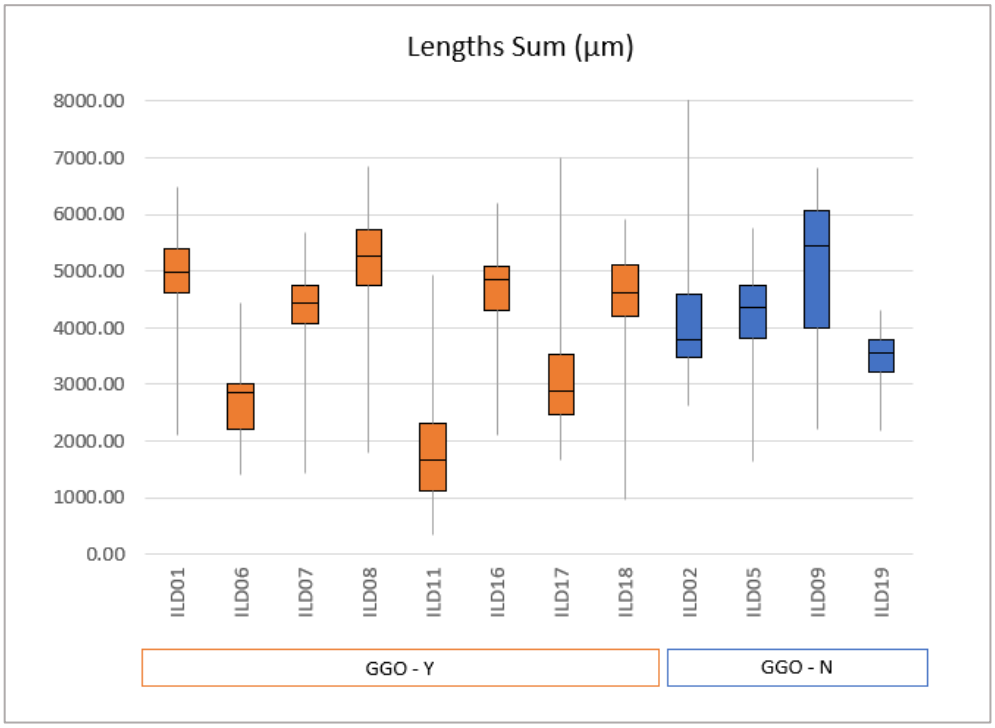

Figure 4.14 – Boxplots representing the Lengths Sum (sum of the lengths of all the detected fibers) for each patient in µm. Patients with ground-glass opacities (GGO) described in their HRCT appear in orange (left), and patients without GGO appear in blue (right).

When looking at the fibrotic patients in Figure 4.15, since the densely fibrotic patients show a smaller median Length but higher Lengths Sum, it seems that these patients have more fibers with a smaller length. However, this argument can be discredited by the fact that sometimes a bigger fiber was misidentified as being multiple smaller fibers

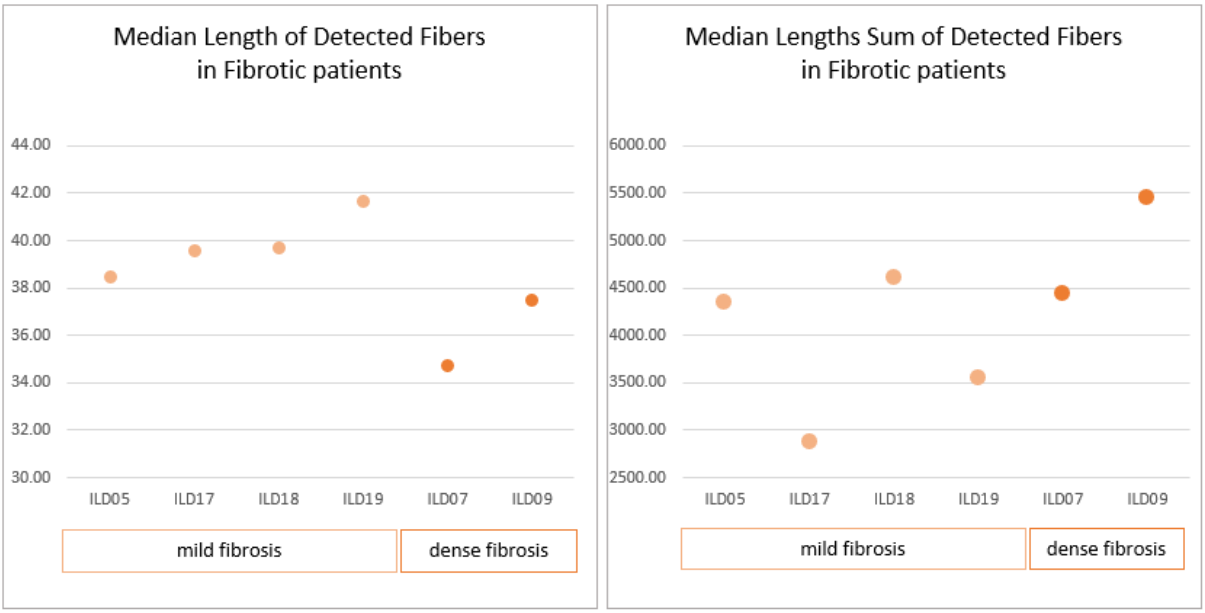

Figure 4.15 – Median values for the Length of detected fibers (left panel) and Lengths Sum (right panel) in the subgroup of patients described as having fibrosis in their biopsies. The mild fibrosis is represented in light orange (left) and dense fibrosis appears darker (right)

Regarding the width of the detected fibers, the mean and standard deviation calculated from the median values represented in Figure 4.16 is $(11.64 \pm 0.3)\mu m$. There seems to be no relevant difference between groups, as seen in Figure 4.16.

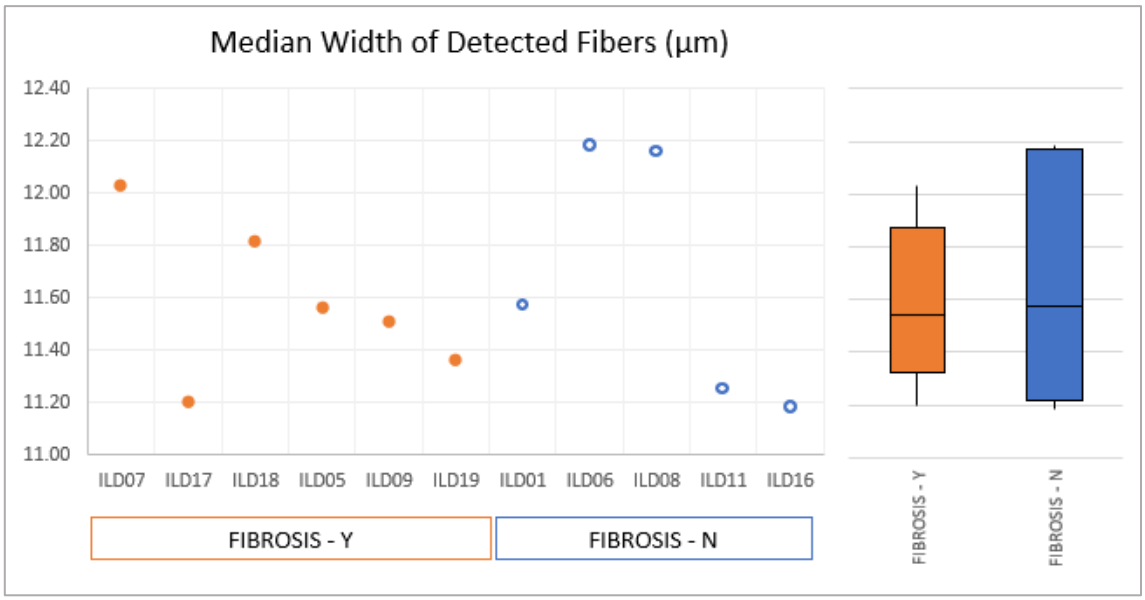

Figure 4.16 – Median values for the width of the detected fibers. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right). The boxplots on the right panel summarize the median Width values for the patient groups relative to the presence of fibrosis.

The following figures (Figures Figure 4.17 to Figure 4.24) show the graphs obtained for the GLCM features results. At the left is represented the median value for each patient and, on the right side, the boxplots summarize the median values from the groups Fibrotic-Y/N. For the variables Entropy and Energy, the plots showing the groups GGO-Y/N are also represented.

It was expected for the Entropy to be higher in patients with fibrosis, and ground-glass opacities, since this feature is associated with the disorder of the image and the fibrotic process is associated with the deposition of disorganized fibers which make the image more complex.

Conversely, the Angular Second Moment (or Energy) was expected to be higher in non-fibrotic patients since their lung tissue should be more homogenous. Also, a high Inverse Difference Moment would be more likely for non-fibrotic patients which contain fewer inhomogeneous areas.

Looking at the difference between patient groups concerning the presence of ground-glass opacities in the Entropy and Energy plots (Figure 4.17 and Figure 4.19), the interquartile range is wider for the GGO-Y group in both features. While the median value of Entropy is higher in the GGO-Y, comparing with GGO-N, the opposite happens for the Energy.

The Contrast, which measures the local variations, was expected to be higher in the fibrotic patients whose lung tissue have more fibers, therefore, edges in the CLE images. However, the contrast also increases with the contributions from noise.

It was expected for the Variance to increase for fibrotic patients since it is also a measure of heterogeneity.

The Correlation measures the linear dependency of grey levels. It seems like the patients with ground-glass opacities on their HRCT have less linear dependencies than the patients without GGO.

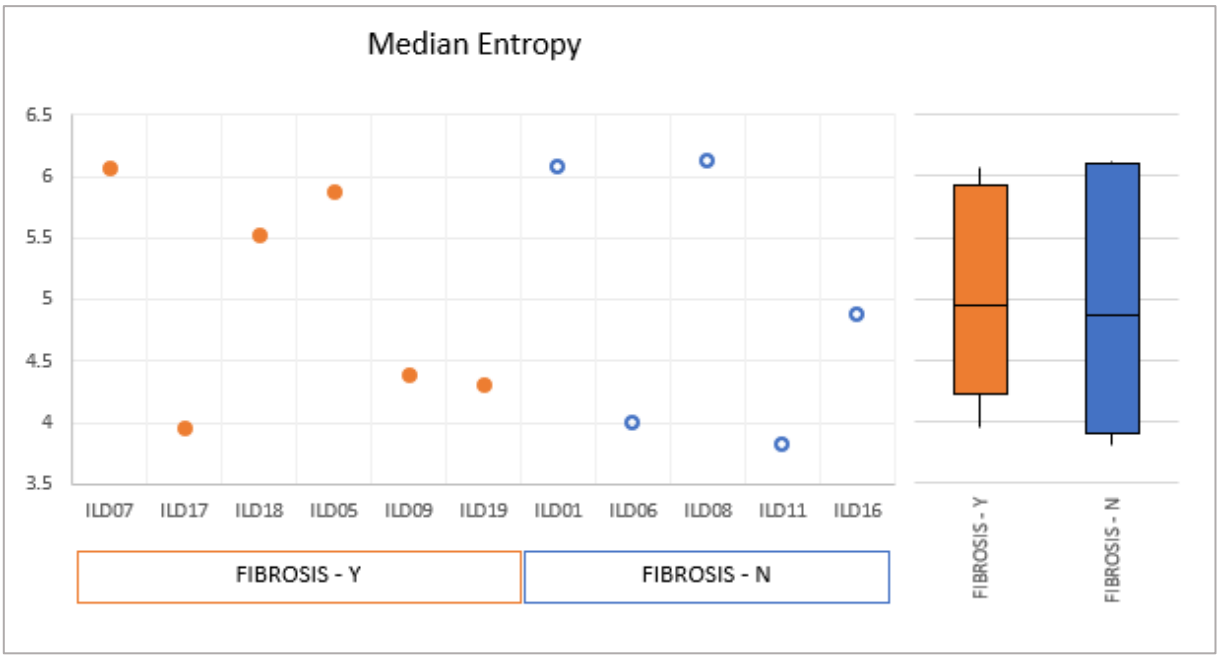

Figure 4.18 – At the left, the plot represents the median values for the Entropy texture feature for each patient. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right). The boxplots on the right summarize the median Entropy values for the patient groups relative to the presence of fibrosis.

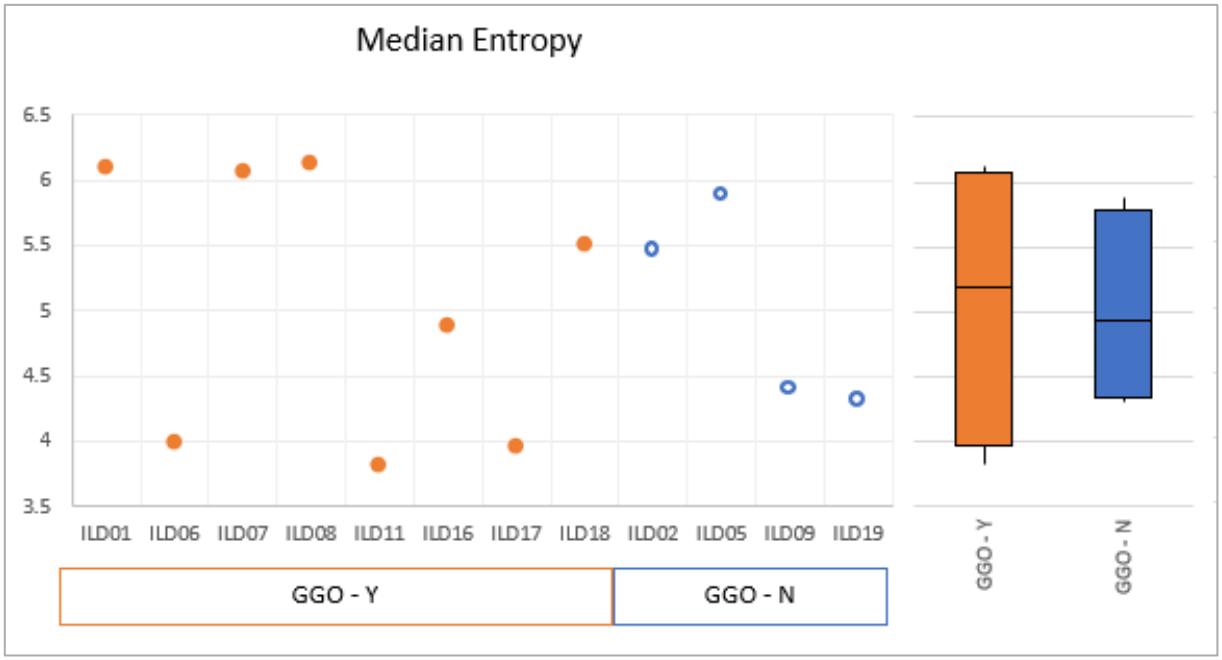

Figure 4.17 – At the left, the plot represents the median values for the Entropy texture feature for each patient. Patients with ground-glass opacities (GGO) in their HRCT appear in orange (left), and patients without GGO appear in blue (right). The boxplots on the right summarize the median Entropy values for the patient groups concerning the presence of GGO.

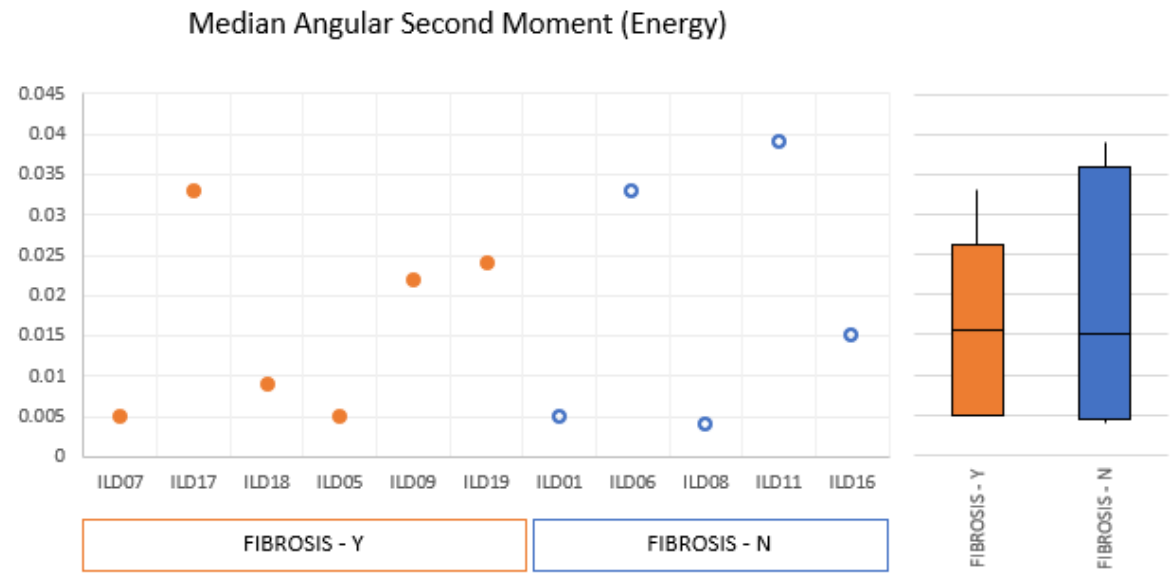

Figure 4.20 – At the left, the plot represents the median values for the Energy texture feature for each patient. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right). The boxplots on the right summarize the median Energy values for the patient groups relative to the presence of fibrosis.

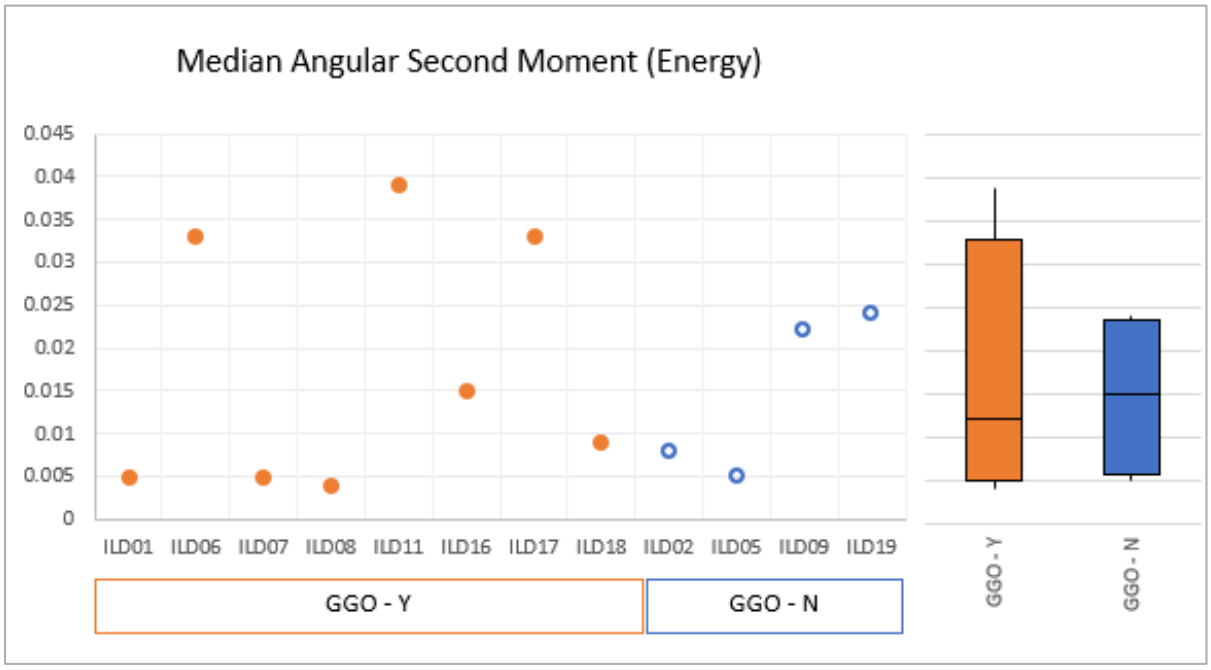

Figure 4.19 – At the left, the plot represents the median values for the Energy texture feature for each patient. Patients with ground-glass opacities (GGO) in their HRCT appear in orange (left), and patients without GGO appear in blue (right). The boxplots on the right summarize the median Energy values for the patient groups concerning the presence of GGO.

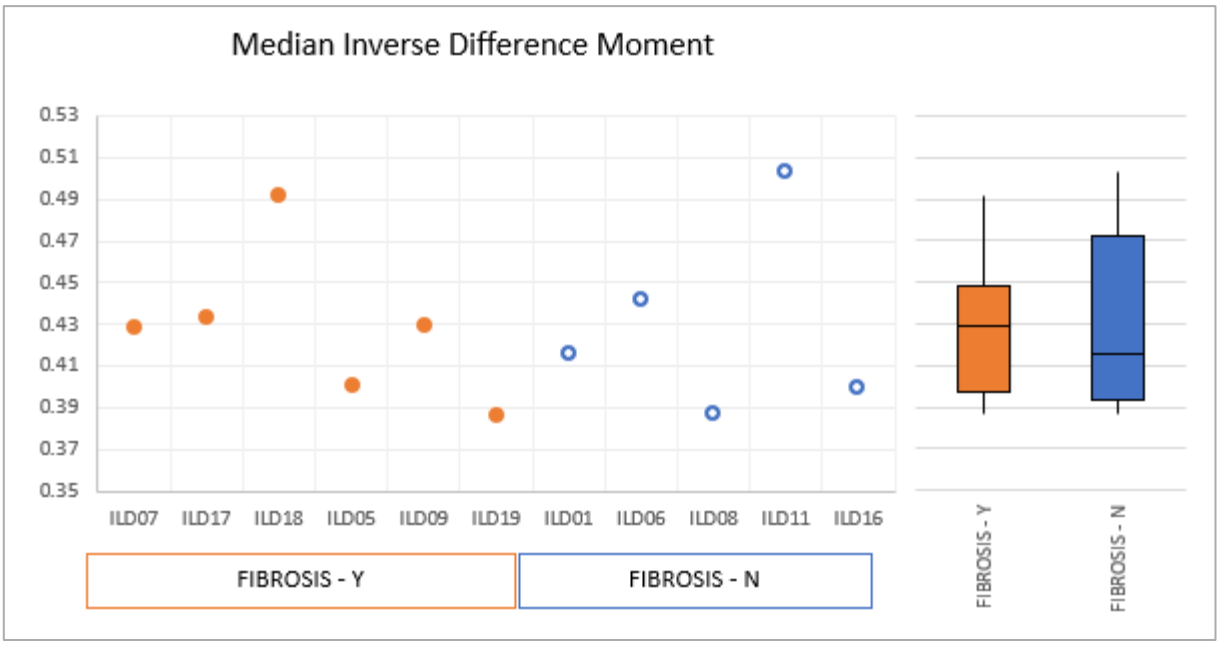

Figure 4.21 – At the left, the plot represents the median values for the Inverse Difference Moment texture feature for each patient. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right). The boxplots on the right summarize the median Inverse Difference Moment values for the patient groups relative to the presence of fibrosis.

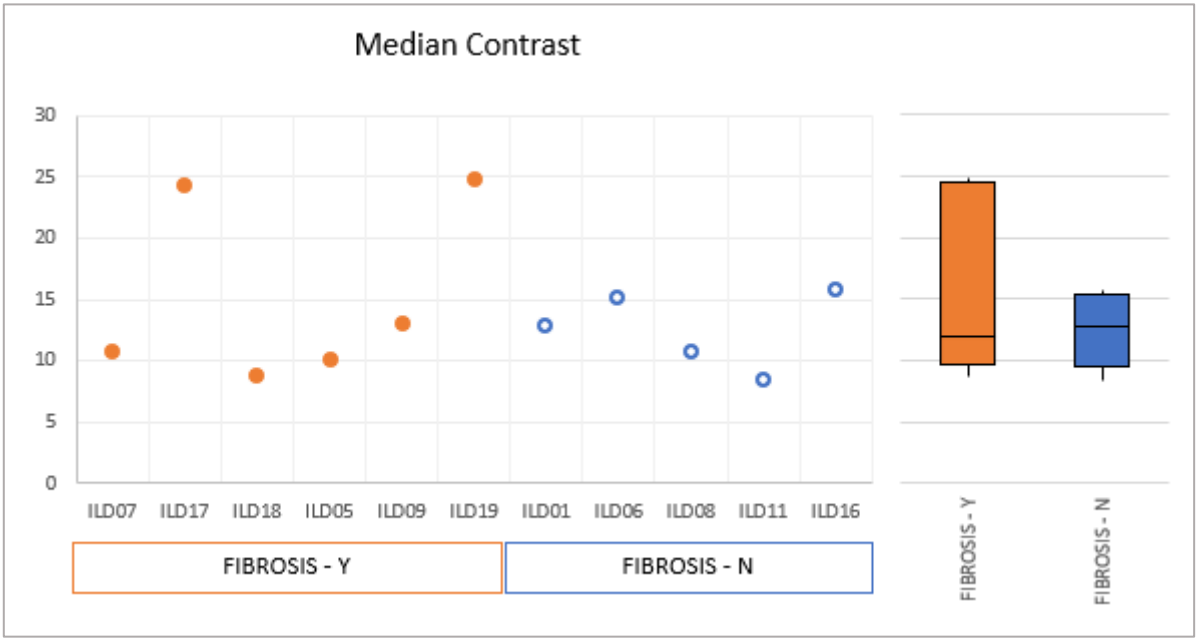

Figure 4.22 – At the left, the plot represents the median values for the Contrast texture feature for each patient. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right). The boxplots on the right summarize the median Contrast values for the patient groups relative to the presence of fibrosis.

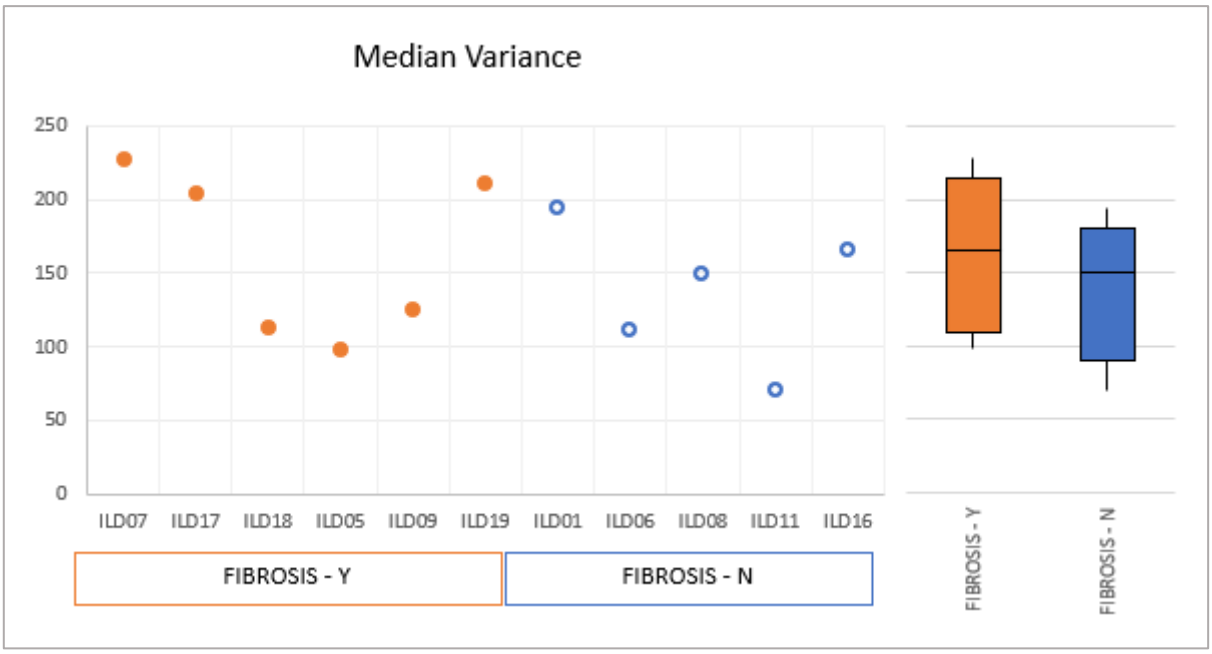

Figure 4.23 – At the left, the plot represents the median values for the Variance texture feature for each patient. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right). The boxplots on the right summarize the median Variance values for the patient groups relative to the presence of fibrosis.

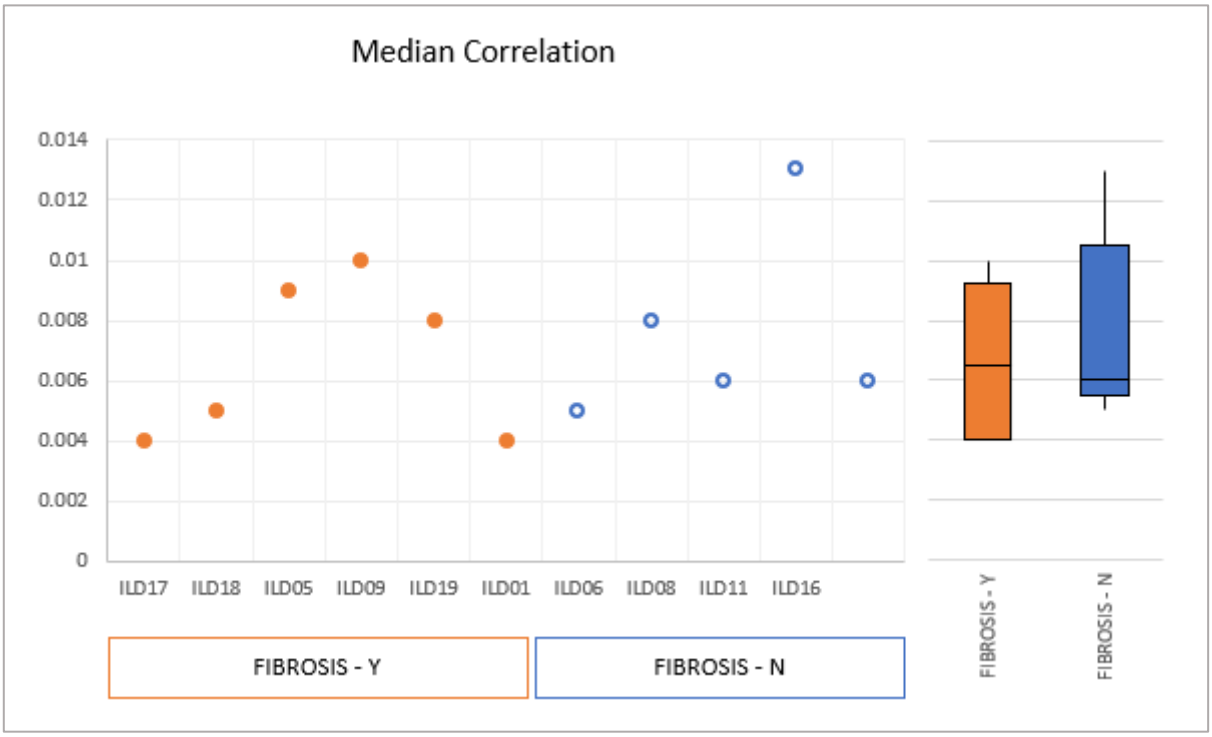

Figure 4.24 – At the left, the plot represents the median values for the Correlation texture feature for each patient. Patients with fibrosis described in their biopsies appear in orange (left), and patients without fibrosis appear in blue (right). The boxplots on the right summarize the median Correlation values for the patient groups relative to the presence of fibrosis.

5 DISCUSSION

The CLE movies imaged from this patient group showed interesting images where elastin fibers are visible as well as inflammatory cells and other structures (vessels, for example). The microscopic resolution adds detail to the HRCT images and the three-dimensionality present adds information regarding the alveolar networks to the histological material. Also, the real-time imaging including the natural breathing movements complements both these techniques. The CLE procedure proved being safe and minimally invasive.

From the original CLE movies, it was possible to generate mosaics from stable movie sequences using the Cellvizio Viewer software, which was already an innovative practice introduced to the handling of these images. The mosaics are very useful since they provide a broader overview of the patients' lung tissue and condense in a static image the information from a movie sequence.

Some movie sequences with low technical quality were excluded from the dataset. Then, the CLE movies were converted into frames which allowed the subsequent image analysis. More than half of the CLE movie sequences presented a noise pattern in the shape of a grid caused by electrical interference which would hamper the following analysis. A "notch filter" was designed to filter in the frequency domain the interference peaks, removing the noise pattern from the images, therefore improving the image quality.

Afterwards, the quantitative analysis enabled to extract objective information from the images in an automatic and unbiased way. It was possible to obtain the Number of Detected Fibers, their Length and Width, the Number of Junctions between fibers and to calculate the Sum from all Fibers' Lengths. Additionally, textural features were extracted: Angular Second Moment (Energy), Entropy, Inverse Difference Moment, Contrast, Variance and Correlation.

The ridge detection algorithm detected most of the elastin fibers recognized by the human eye. Sometimes, some fibers were not detected, probably due to poor contrast or image quality, or a fiber was detected as multiple smaller fibers. Even with some detection errors, this algorithm is more effective and efficient than a human observer since it is more consistent within images and it processes the entire dataset in a few seconds. These measurements related to the elastin fibers can be valuable to identify and follow fibrotic patients.

Regarding the texture analysis, the GLCM algorithm enabled to assign objective values to the texture present in the images. This approach enables to overcome the bias associated with human visual scoring. It is also a much efficient process, saving time from the clinicians which would have to watch the movies frame-by-frame and score them. This algorithm can compute several texture features simultaneously, which can later be selected. The Haralick Features were chosen to be applied in this project, but other texture features such as the Local Binary Patterns could also be tested.

It would be intuitive to analyze the pixel intensity histograms of the CLE images, because more fluorescence is associated with more elastin and other fluorescent structures. However, the MKT company which manufactures the CLE instrument, whose CEO and team kindly met us in their head office in Paris, warned us regarding the laser automatic control - an algorithm in the Cellvizio machine that adjusts the intensity of the image during the acquisition according to the collected intensity. Since this algorithm is by default enabled so the clinicians get real-time images with corrected contrast, all the data acquired for this study was also adjusted by this. This process is not reversible. In this sense, they advised us to perform morphology studies instead. Besides studying the morphology, we opted to use the GLCM because it relies on relationships between pixel intensities and not in the pixel value itself.

In a future study, where quantitative analysis is prioritized over visual quality, it should be asked to the MKT company to disable the laser automatic control before the data is acquired.

Regarding the image dataset, some images were easily compared to the reference of healthy alveolar tissue, and other looked very different from that, not meaning these were obviously unhealthy. The quality of the images varies within and between patients with some having good contrast and others poor contrast. Multiple movie sequences were excluded from the dataset mainly because few alveolar structures were visible, the image sharpness was considered unacceptable, or the tissue was stretched (due to the probe), or distorted (due to movement artifacts). It is hard to know if the poor contrast and lack of alveolar structures is associated with the acquisition of the data – over or under pulling the probe – or with the disease state itself. As an example, some of these patient's histopathological description mentioned the presence of alveolar filling, which probably diminishes the visibility of the alveolar structures. Also, some ILD are associated with destruction of alveolar network. The fact that this was the first time the research group used this technology can justify some of the difficulties in obtaining good quality images.

For the above reasons, the data selection based on the images technical quality can be unfair. Particularly the criterion a) "Lack of alveolar structures. Specifically, less than 75% of the FOV", since this absence can sometimes be associated with the disease state, being a discriminating factor, which would be useful for the quantitative analysis. Also, some bias is involved in this process since it was performed by a group of two human observers. However, similar selections were performed in other studies and they guarantee less images to be processed and a smaller dataset for visual analysis.

Because ILDs are diffuse and different patients might have different disease distribution and state, the location and extent of the biopsy or CLE imaging will influence the diagnosis. If a pathologist sees a biopsy having parts with normal lung tissue and others with severely diseased tissue, he can decide – according to the severity and distribution – how to weight the contributions from healthy and unhealthy tissue. With the CLE sequences that is harder to do: if the movie sequence spent more seconds in the healthy tissue and less in the diseased tissue, that will have more weight over the quantitative analysis because more frames with diseased tissue will be processed by the algorithms. If mosaics were used in the image analysis, this issue could be diminished. The choice of the statistical variables used – in this study the median (and minimum, maximum, 1st and 3rd quartiles) – also affects the weight that the contribution from each frame will have on the analysis.

It is relevant to note that as the probe is being pushed further, more planes with more fibers will be shown on the resultant CLE image.

The few CLE images performed *ex vivo* show better quality probably because the specimen is submerged in saline solution which might wash some fluids such as sputum which hamper the image quality. However, the presence of this fluid can be discriminative of an inflammatory state in the image texture. Additionally, the clinician has more control over the probe which gives more stable videos with more sharp images. Although the mosaics are useful to visually analyze the alveolar network, they were not used in the quantitative analysis since the irregular shape of the FOV makes it difficult to apply the GLCM algorithm. In order to apply the other measurements, the FOV area should be calculated to normalize the values.

Regarding the results from the quantitative analysis, they were plotted in multiple ways and analyzed. To analyze the results, the outcome from the radiologic and histopathological descriptions was used as a reference. However, to compare CLE with the existing diagnosing techniques is a difficult process because they differ in terms of scale and size of the field-of-view, with the histopathology being associated with a micro scale like the CLE but having a broader field-of-view which allows a quick

overview of the lung tissue state. In terms of the outcome from these imaging techniques, while in histopathology the clinician is mostly focused on the collagen fibers, in CLE the visible structures are almost exclusively elastin fibers.

As mentioned, the measurements related to the elastin fibers intend to help identifying the fibrotic patients. It was expected from the fibers from fibrotic patients to be wider, but that was not observed. Also, it was expected that this group of patients showed more fibers and junctions, but there was no significant difference between groups. However, when the fibrotic group is segregated, the number of fibers and junctions seems to separate the mild from the dense fibrosis. This is a very interesting result because it suggests that monitoring the number of fibers/junctions with CLE could potentially be used as anti-fibrotic medication efficacy measuring tool. Currently, the monitoring of this medication efficacy is performed using CT-scan images and lung function tests. The CLE could potentially reveal the antifibrotic treatment-effects earlier.

With respect to the texture measurements, it was expected for the fibrotic patients to have higher values of Entropy, Contrast and Variance and low Energy and Inverse Difference Moment since their lung tissue should correspond to more complex, heterogeneous images with more edges present. Still, no significant difference between groups can be stated yet.

Although the results with the present data set didn't demonstrate a strong correlation between the CLE and HRCT/histopathology outcomes, the measurements themselves already contribute to the study of Interstitial Lung Diseases, aiding to the knowledge of the disease physiology. In fact, the patient group recruited for this study was limited by the maximum number of procedures allowed by the CLE mini probe – 20. The sample size got even smaller because some patients were excluded. With a bigger dataset, it is expected that some measurements will correlate with the other diagnosing techniques and enable to stratify the patients into groups and apply data classification. It is possible to speculate that some of the studied variables are better candidates to be used in the classification of patients, namely, the Number of Fibers and Junctions, the Sums Length, and the Haralick features Entropy and Energy.

6 Conclusion and Future Work

This project approached the clinical problems associated with the diagnosis of Interstitial Lung Diseases (ILD). ILD is a heterogenous group of more than 200 diseases which requires a multidisciplinary discussion (MDD) to be diagnosed. Although the expertise associated with the multidisciplinary teams, 10% of the ILD patients are categorized as unclassifiable, mostly due to the absence of histopathological data justified by the risks of taking biopsies to these patients.

To determine the specific ILD a patient suffers from is very important because it influences the treatment and management, being crucial to identify the fibrotic patients, particularly the Idiopathic Pulmonary Fibrosis (IPF), which have a more restricted and expensive treatment and usually a worse prognosis.

The Confocal Laser Endomicroscopy can add relevant information to the MDD through a minimally invasive bronchoscopic procedure. The CLE showed very interesting images with microscopic resolution where the elastin fibers are visible through their autofluorescence. There is evidence that the amount and architecture of the elastin fibers is altered in diseased states, particularly in fibrotic lung tissue.

Therefore, we performed some measurements related to the amount and size of the elastin fibers detected in the CLE images. These measurements can potentially identify the fibrotic patients or function as a minimally invasive anti-fibrotic medication efficacy tool. Indeed, with our results, we can expect the CLE to differentiate between the mild and dense fibrotic state.

Image textures are complex visual patterns composed of entities or regions with sub-patterns with the characteristics of brightness, color, shape, size, etc. Some disease states can be revealed by textural changes in medical images detected by a computer. In some cases, these changes are hard even for a trained viewer to recognize specially in early disease stages. In fact, the diseased lung tissue was visually different from the healthy tissue in the CLE images. Thus, we extracted texture values which would reflect this difference in a consistent way. We expected the texture features associated with complexity and heterogeneity to arouse higher values in fibrotic patients.

This project allowed the extraction of numerical features from the CLE movies, which is preferred than to simply visually analyzing these images, avoiding unbiased and time-consuming visual scorings. Our input was a set of interesting CLE movies, which we preprocessed, producing mosaics and enhancing the image quality, and converted into frames for the quantitative analysis. Our outcome became a set of objective and consistent values describing the image, as well as mosaics and filtered images which should also contribute for the clinician analysis.

By imaging with CLE one patient's biopsy which contained pleura (associated with a pneumothorax) confirmed by the pathologist, it was proved that CLE images allow the distinction between the elastin fibers from pleura – which are arranged in a different way than the fibers from the alveoli or alveolar ducts. This finding was described in a brief communication entitled "Confocal Laser Endomicroscopy (CLE) as a Guidance Tool for Transbronchial Lung Cryobiopsies", L. Wijmans et al.[73], presented at the European Respiratory Society (ERS) International Congress 2018 (17 September 2018) and published at the Respiration Journal. In Appendix II, images from this publication are presented.

Although our sample size was very low, the problem addressed in this project affects hundreds of people who suffer from Interstitial Lung Diseases but have an uncertain diagnosis. Our project intends to give more information to the clinicians which make up the multidisciplinary team, aiding the decision process, particularly in the cases when the patient cannot undergo an invasive procedure such as a lung

biopsy. Due to the small number of patients included in this study, the results obtained cannot discriminate between groups yet. Nevertheless, the measurements themselves already contribute to the study of ILD, helping to understand the disease physiology.

Further research is still needed so this technique can be used in clinical practice. In the future, it would be desirable to apply the measurements from this project to a larger patient cohort. Also, it would be interesting to try other texture features, such as Local Binary Patterns (LBP). With the automatic laser control disabled, it would be possible to perform intensity related measurements. Afterwards, data classification could be applied. I would suggest the use of support vector machine with recursive feature elimination (SVM-RFE): using the classification from the histopathological description (classes fibrotic and non-fibrotic), through supervised learning, a model can be built to assign new examples – which weren't in the training set – to one class or the other. The recursive feature elimination weights the different attributes (in this case, the several morphometric measurements and texture features) and generates the best combination of attributes that will give the best prediction on the class. The added value of the CLE could then be compared to the existing techniques through a Receiver Operating Characteristic (ROC) curve which plots the true positive rate (sensitivity) against the false positive rate (1 – specificity) at various threshold settings. This ROC could evaluate the performance of the clinicians when aided by the CLE data, comparing for example, their opinion when analyzing the CLE data before knowing the histopathological results, with their opinion in the presence of histopathological data.

The main goal of this project to improve the CLE technique and increase its usability was achieved. The extraction of numerical values expressing the state of the alveolar space already contributes to the study of the Interstitial Lung Diseases and, when applied to a larger sample, will potentially allow the stratification of the ILD patients. Eventually, in the future, CLE will help reducing the need for lung biopsies.

7 APPENDIX I

Examples of Confocal Laser Endomicroscopy movies in ILD patients

The following url leads to a cloud storage containing two examples of Confocal Laser Endomicroscopy movies from patients with Interstitial Lung Diseases: one imaged *in vivo* and the other imaged *ex vivo*. The videos speed has been reduced to improve the visualization.

https://www.dropbox.com/sh/5rvipum20uhh5jl/AABuVNn_JYakBuJUm6I8QkWNa?dl=0

8 APPENDIX II

Images from the publication "Confocal laser endomicroscopy (CLE) as a guidance tool for transbronchial lung cryobiopsies in ILD", L. Wijmans, et al. [73], presented at the European Respiratory Society (ERS) International Congress 2018 (17 September 2018) and accepted for publication at the Respiration Journal.

To date, guidance of transbronchial cryobiopsies (TBCB) is performed by fluoroscopy: being 2D hampers the identification of pleura/diaphragm-distance; its resolution complicates real-time identification of fibrotic areas; the radiation dose is a hazard for both patient and the medical team.

Since real-time CLE was capable of identifying key-characteristics that affect the diagnostic yield (dense fibrotic areas and subpleural space) and complication rate during TBCB (pleura and subpleural space), this technique could be used as a guidance tool during TBCB-procedures. Indeed, the concept of a smart-cryoprobe to minimize complication and non-diagnostic rate of TBCB in ILD patients was proposed in this publication.

Figure 8.1 shows the CLE imaging where the pleura – with a dense cross-fiber pattern – and the adjacent alveoli of the subpleural space are visualized. This figure also shows the corresponding histology image which was stained for elastin with EVG showing the elastin fibers in thin dark lines.

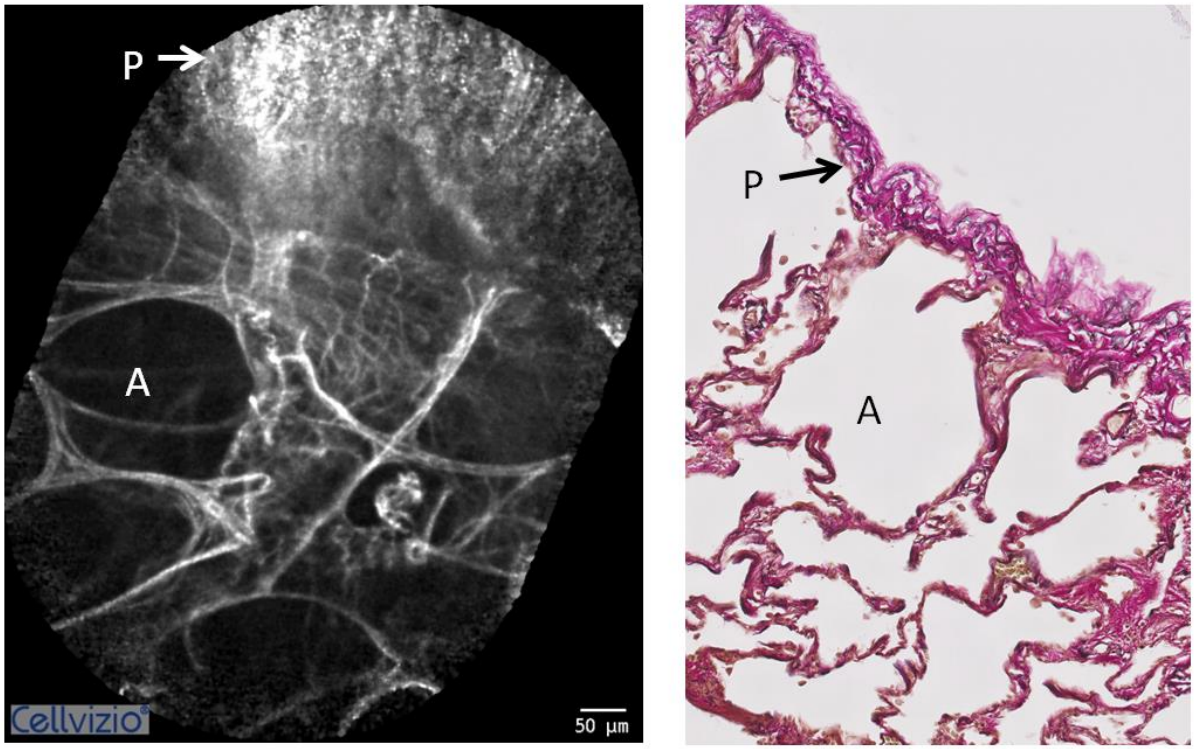

Figure 8.1 – Ex-vivo CLE imaging visualizes pleura (P) and the adjacent subpleural space with alveolar airspaces (A) and elastin fibers of the alveolar septa (S) (left panel) and in histology image from corresponding location (right panel). This patient was diagnosed with a hypersensitivity pneumonitis. The thin dark lines in the histology image are the elastin fibers (stained with EVG). Image from [73]

9 REFERENCES

[1] S. International and M. Consensus, "American Thoracic Society / European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias," vol. 165, pp. 277–304, 2002.

[2] C. J. Ryerson *et al.*, "Prevalence and prognosis of unclassifiable interstitial lung disease," *Eur. Respir. J.*, vol. 42, no. 3, pp. 750–757, 2013.

[3] H. E. Jo *et al.*, "Clinical impact of the interstitial lung disease multidisciplinary service," *Respirology*, vol. 21, no. 8, pp. 1438–1444, 2016.

[4] S. A. Guler, S. A. Berezowska, A. Christe, T. Geiser, and M. Funke-Chambour, "Multidisciplinary discussion for diagnosis of interstitial lung disease in real life," *Swiss Med. Wkly.*, vol. 146, no. June, pp. 1–4, 2016.

[5] G. Raghu *et al.*, "An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline," *Am. J. Respir. Crit. Care Med.*, vol. 192, no. 2, pp. e3–e19, 2015.

[6] American Thoracic Society, "Interstitial Lung Disease," in *Breathing in America: Diseases, Progress, and Hope*, D. E. Schraufnagel, Ed. 2010, pp. 99–108.

[7] The Idiopathic Pulmonary Fibrosis Clinical Research Network, "Prednisone, Azathioprine, and N-Acetylcysteine for Pulmonary Fibrosis," *N. Engl. J. Med.*, vol. 366, no. 21, pp. 1968–1977, 2012.

[8] T. R. Harrison, A. S Fauci, D. L Kasper, and D. L Longo, "Interstitial Lung Diseases," in *Harrison`s Principles of Internal Medicine*, 2012, pp. 2160–2170.

[9] D. B. Coultas, R. E. Zumwalt, W. C. Black, and R. E. Sobonya, "The epidemiology of interstitial lung diseases.," *Am. J. Respir. Crit. Care Med.*, vol. 150, no. 4, pp. 967–972, 1994.

[10] J. Gribbin, R. B. Hubbard, I. Le Jeune, C. J. P. Smith, J. West, and L. J. Tata, "Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK," *Thorax*, vol. 61, no. 11, pp. 980–985, 2006.

[11] G. Raghu, D. Weycker, J. Edelsberg, W. Z. Bradford, and G. Oster, "Incidence and prevalence of idiopathic pulmonary fibrosis," *Am. J. Respir. Crit. Care Med.*, vol. 174, no. 7, pp. 810–816, 2006.

[12] U. Wells, "Histopathologic Diagnosis in Diffuse Lung Disease An Ailing Gold Standard," *Am. J. Respir. Crit. Care Med.*, vol. 170, no. 8, pp. 827–828, 2004.

[13] G. Raghu *et al.*, "An Official ATS/ERS/JRS/ALAT Statement: Idiopathic pulmonary fibrosis: Evidence-based guidelines for diagnosis and management," *Am. J. Respir. Crit. Care Med.*, vol. 183, no. 6, pp. 788–824, 2011.

[14] T. a Wynn, V. G. Yugandhar, and M. a. Clark, "Cellular and molecular mechanisms of fibrosis," *J Pathol*, vol. 46, no. 2, pp. 26–32, 2013.

[15] R. M. Cherniack *et al.*, "Quantitative assessment of lung pathology in idiopathic pulmonary fibrosis," *Am Rev Respir Dis*, vol. 144, no. 4, pp. 892–900, 1991.

[16] R. G. Crystal, P. B. Bitterman, S. I. Rennard, A. J. Hance, and B. A. Keogh, "Interstitial Lung Diseases of Unknown Cause. Disorders characterized by chronic inflammation of the lower respiratory tract. (First of two parts)," *N. Engl. J. Med.*, vol. 310, no. 3, 1984.

[17] R. G. Crystal, P. B. Bitterman, and S. I. Rennard, "Interstitial lung diseases of unknown cause. Disorders characterized by chronic inflammation of the lower respiratory tract. (Second of two parts)," *N. Engl. J. Med.*, vol. 310, no. 4, pp. 235–244, 1984.

[18] U. Specks, A. Nerlich, T. V Colby, I. Wiest, and R. Timpl, “Increased expression of type VI collagen in lung fibrosis,” *Am J Respir Crit Care Med*, vol. 151, no. 6, pp. 1956–1964, 1995.

[19] M. Toshima, Y. Ohtani, and O. Ohtani, “Three-dimensional architecture of elastin and collagen fiber networks in the human and rat lung,” *Archives of Histology and Cytology*, vol. 67, no. 1. pp. 31–40, 2004.

[20] T. J. Mariani, E. Crouch, J. D. Roby, B. Starcher, and R. A. Pierce, “Increased elastin production in experimental granulomatous lung disease,” *Am.J.Pathol.*, vol. 147, no. 4, pp. 988–1000, 1995.

[21] S. Radha, N. Ravindra, T. Afroz, and S. Prasad, “Diagnostic utility of bronchoalveolar lavage,” *J. Cytol.*, vol. 31, no. 3, p. 136, 2014.

[22] K. C. Meyer *et al.*, “An official American Thoracic Society clinical practice guideline: The clinical utility of bronchoalveolar lavage cellular analysis in interstitial lung disease,” *Am. J. Respir. Crit. Care Med.*, vol. 185, no. 9, pp. 1004–1014, 2012.

[23] K. C. Meyer and G. Raghu, “Bronchoalveolar lavage for the evaluation of interstitial lung disease: Is it clinically useful?,” *Eur. Respir. J.*, vol. 38, no. 4, pp. 761–769, 2011.

[24] G. Yang, C. Sau, W. Lai, J. Cichon, and W. Li, “Comparison of Transbronchial and Cryobiopsies in Evaluation of Diffuse Parenchymal Lung Disease,” *J. Bronchology Interv. Pulmonol.*, vol. 23, no. 1, pp. 1173–1178, 2016.

[25] G. L. Casoni *et al.*, “Transbronchial lung cryobiopsy in the diagnosis of fibrotic interstitial lung diseases,” *PLoS One*, vol. 9, no. 2, pp. 1–7, 2014.

[26] S. Tomassetti *et al.*, “Bronchoscopic lung cryobiopsy increases diagnostic confidence in the multidisciplinary diagnosis of idiopathic pulmonary fibrosis,” *Am. J. Respir. Crit. Care Med.*, vol. 193, no. 7, pp. 745–752, 2016.

[27] M. E. Kreider *et al.*, “Complications of Video-Assisted Thoracoscopic Lung Biopsy in Patients with Interstitial Lung Disease,” *Ann. Thorac. Surg.*, vol. 83, no. 3, pp. 1140–1144, 2007.

[28] T. E. King *et al.*, “A Phase 3 Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis,” *N. Engl. J. Med.*, vol. 370, no. 22, pp. 2083–2092, 2014.

[29] K. O. Leslie, “My approach to interstitial lung disease using clinical, radiological and histopathological patterns.,” *J. Clin. Pathol.*, vol. 62, no. 1, pp. 387–401, 2009.

[30] N. Sverzellati, D. A. Lynch, D. M. Hansell, T. Johkoh, T. E. King, and W. D. Travis, “American Thoracic Society–European Respiratory Society Classification of the Idiopathic Interstitial Pneumonias: Advances in Knowledge since 2002,” *RadioGraphics*, vol. 35, no. 7, pp. 1849–1871, 2015.

[31] R. C. Henrikson, J. E. Mazurkiewicz, and G. I. Kaye, “Elastic Fibers,” in *Histology*, 1st ed., Baltimore: Williams & Wilkins, 1997.

[32] A. Lichtenberg, S. Cebotari, I. Tudorache, A. Hilfiker, and A. Haverich, “Biological Scaffolds for Heart Valve Tissue Engineering,” in *Tissue Engineering*, Second., Hauser, Hansjörg, Fussenegger, and M. M., Eds. Methods in Moleculer Medicine, 2007, p. 313.

[33] G. F. Rozin, M. M. Gomes, E. R. Parra, R. A. Kairalla, C. R. R. De Carvalho, and V. L. Capelozzi, “Collagen and elastic system in the remodelling process of major types of idiopathic interstitial pneumonias (IIP),” *Histopathology*, vol. 46, no. 4, pp. 413–421, 2005.

[34] N. Enomoto *et al.*, “Quantitative analysis of lung elastic fibers in idiopathic pleuroparenchymal fibroelastosis (IPPFE): Comparison of clinical, radiological, and pathological findings with those of idiopathic pulmonary fibrosis (IPF),” *BMC Pulm. Med.*, vol. 14, no. 1, pp. 1–10, 2014.

[35] N. Enomoto *et al.*, “Amount of elastic fibers predicts prognosis of idiopathic pulmonary fibrosis,” *Respir. Med.*, vol. 107, no. 10, pp. 1608–1616, 2013.

[36] E. M. Negri, G. S. Montes, P. H. N. Saldiva, and V. L. Capelozzi, “Architectural remodelling in

acute and chronic interstitial lung disease: Fibrosis or fibroelastosis?," *Histopathology*, vol. 37, no. 5, pp. 393–401, 2000.

[37] "Confocal Miniprobe Range." [Online]. Available: http://www.cellvizio.net/medical-library/educational-materials. [Accessed: 02-Sep-2017].

[38] S. S. Chauhan *et al.*, "Confocal laser endomicroscopy," *Gastrointest. Endosc.*, vol. 80, no. 6, pp. 928–938, 2014.

[39] J. M. Jabbour, M. A. Saldua, J. N. Bixler, and K. C. Maitland, "Confocal Endocmicroscopy: Instrumentation and Medical Applications," *Ann Biomed Eng*, vol. 40, no. 2, pp. 378–397, 2012.

[40] O. Danilevskaya, A. Averyanov, V. Lesnyak, A. Chernyaev, and A. Sorokina, "Confocal laser endomicroscopy for diagnosis and monitoring of pulmonary alveolar proteinosis," *J Bronchol. Interv Pulmonol*, vol. 22, no. 1, pp. 33–40, 2015.

[41] L. Thiberville, S. Moreno-Swirc, T. Vercauteren, E. Peltier, C. Cavé, and G. B. Heckly, "In Vivo imaging of the bronchial wall microstructure using fibered confocal fluorescence microscopy," *Am. J. Respir. Crit. Care Med.*, vol. 175, no. 1, pp. 22–31, 2007.

[42] L. Thiberville *et al.*, "Confocal fluorescence endomicroscopy of the human airways.," *Proc. Am. Thorac. Soc.*, vol. 6, no. 5, pp. 444–9, 2009.

[43] L. Thiberville *et al.*, "Human in vivo fluorescence microimaging of the alveolar ducts and sacs during bronchoscopy," *Eur. Respir. J.*, vol. 33, no. 5, pp. 974–985, 2009.

[44] L. Thiberville and M. Salaün, "Bronchoscopic advances: on the way to the cells.," *Respiration.*, vol. 79, no. 6, pp. 441–449, 2010.

[45] "Healthy Alveoli CLE Video." [Online]. Available: www.cellvizio.net/self-training. [Accessed: 01-Mar-2017].

[46] M. Salaün and F. Roussel, "In vivo imaging of pulmonary alveolar proteinosis using confocal endomicroscopy," *Eur. Respir. J.*, vol. 36, no. 2, pp. 451–453, 2010.

[47] J. Yserbyt, C. Dooms, V. Ninane, M. Decramer, and G. Verleden, "Perspectives using probe-based confocal laser endomicroscopy of the respiratory tract," *Swiss Med. Wkly.*, vol. 143, no. March, pp. 1–11, 2013.

[48] O. Danilevskaya, A. Averyanov, V. Lesnyak, A. Chernyaev, and A. Sorokina, "Confocal laser endomicroscopy for diagnosis and monitoring of pulmonary alveolar proteinosis.," *J. Bronchology Interv. Pulmonol.*, vol. 22, no. 1, pp. 33–40, 2015.

[49] N. A. Bhatt, S. C. Parrish, P. Malafronte, and R. F. Browning, "Characterization Of Pleura By Probe-Based Confocal Laser Endomicroscopy," *Am Thorac. Soc*, 2015.

[50] R. C. Newton, S. V. Kemp, G. Z. Yang, D. S. Elson, A. Darzi, and P. L. Shah, "Imaging parenchymal lung diseases with confocal endomicroscopy," *Respir. Med.*, vol. 106, no. 1, pp. 127–137, 2012.

[51] R. Kiesslich *et al.*, "Confocal laser endoscopy for diagnosing intraepithelial neoplasias and colorectal cancer in vivo," *Gastroenterology*, vol. 127, no. 3, pp. 706–713, 2004.

[52] L. Dalar *et al.*, "Probe-based Confocal Laser Endomicroscopy (pCLE) in the Diagnosis of Diffuse Parenchymal Lung Diseases : Two Cases," *Austin J. Pulm. Respir. Med.*, vol. 1, no. 4, p. 5, 2014.

[53] T. R. Shulimzon, "Real-time vision of a sarcoid granuloma at bronchoscopy," *Am. J. Respir. Crit. Care Med.*, vol. 187, no. 7, p. 776, 2013.

[54] O. Danilevskaya *et al.*, "The case of diagnostics of invasive pulmonary aspergillosis by in vivo probe-based confocal laser endomicroscopy of central and distal airways," *Med. Mycol. Case Rep.*, vol. 5, no. 1, pp. 35–39, 2014.

[55] F. Reichenberger, E. Silbernagel, W. Gesierich, M. Lindner, A. Morresi-Hauff, and J. Behr,

"Imaging pattern in interstitial lung diseases using probe based confocal laser endomicroscopy (pCLE)," *Eur. Respir. J.*, 2015.

[56] P. Meng, G. L. Tan, S. Y. Low, A. Takano, Y. L. Ng, and D. Anantham, "Fibred confocal fluorescence microscopy in the diagnosis of interstitial lung diseases," *J. Thorac. Dis.*, vol. 8, no. 12, pp. 3505–3514, 2016.

[57] J. Yserbyt, T. Alamé, C. Dooms, and V. Ninane, "Pulmonary alveolar microlithiasis and probe-based confocal laser endomicroscopy," *J. Bronchol. Interv. Pulmonol.*, vol. 20, no. 2, pp. 159–163, 2013.

[58] J. Yserbyt, C. Dooms, W. Janssens, and G. M. Verleden, "Supplementary material -Endoscopic advanced imaging of the respiratory tract : exploring probe-based confocal laser endomicroscopy in emphysema," *Thorax*, 2017.

[59] A. S. Wellikoff, R. C. Holladay, G. H. Downie, C. S. Chaudoir, L. Brandi, and E. A. Turbat-Herrera, "Comparison of in vivo probe-based confocal laser endomicroscopy with histopathology in lung cancer: A move toward optical biopsy," *Respirology*, vol. 20, no. 6, pp. 967–974, 2015.

[60] J. Yserbyt, C. Dooms, M. Decramer, and G. M. Verleden, "Probe-based confocal laser endomicroscopy of the respiratory tract: A data consistency analysis," *Respir. Med.*, vol. 107, no. 8, pp. 1234–1240, 2013.

[61] J. Yserbyt, C. Dooms, W. Janssens, and G. M. Verleden, "Endoscopic advanced imaging of the respiratory tract : exploring probe-based confocal laser endomicroscopy in emphysema," *Thorax*, vol. 0, no. 0, pp. 1–4, 2017.

[62] B. André, T. Vercauteren, A. M. Buchner, M. B. Wallace, and N. Ayache, "A smart atlas for endomicroscopy using automated video retrieval," *Med. Image Anal.*, vol. 15, no. 4, pp. 460–476, 2011.

[63] B. André, T. Vercauteren, A. M. Buchner, M. Krishna, N. Ayache, and M. B. Wallac, "Software for automated classification of probe-based confocal laser endomicroscopy videos of colorectal polyps," *World J. Gastroenterol.*, vol. 18, no. 39, pp. 5560–5569, 2012.

[64] C. Désir, C. Petitjean, L. Heutte, L. Thiberville, and M. Salaün, "An SVM-based distal lung image classification using texture descriptors," *Comput. Med. Imaging Graph.*, vol. 36, no. 4, pp. 264–270, 2012.

[65] A. Rakotomamonjy, C. Petitjean, M. Salaun, and L. Thiberville, "Scattering features for lung cancer detection in fibered confocal fluorescence microscopy images.," *Artif. Intell. Med.*, vol. 61, no. 2, pp. 105–118, 2014.

[66] J. T. Annema, "Clinical Trial description: Confocal Laser Endomicroscopy and Optical Coherence Tomography for Diagnosing ILD." [Online]. Available: https://clinicaltrials.gov/ct2/show/NCT02689102. [Accessed: 20-May-2017].

[67] I. The MathWorks, "MATLAB and Image Processing Toolbox Release 2017b," *MATLAB version 9.3.0.713579 (R2017b)*. Natick, Massachusetts, United States, 2017.

[68] U. of R. Department of Computer Science, "Frequency Domain Processing Lab." [Online]. Available: http://www.cs.uregina.ca/Links/class-info/425/Lab5/index.html. [Accessed: 24-Feb-2017].

[69] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 20(2), pp. 113–125, 1998.

[70] RStudio Team, "RStudio: Integrated Development for R. RStudio, Inc." Boston, 2015.

[71] R. M. Haralick, K. Shanmugam, and I. T. S. H. a K. Dinstein, "Textural Features for Image Classification," *IEEE Trans. Syst. Man. Cybern.*, vol. 3, no. 6, 1973.

[72] F. Albregtsen, "Statistical Texture Measures Computed from Gray Level Coocurrence

Matrices," *Bound. 2*, vol. 3, no. 1, p. 45, 2005.

[73] L. Wijmans *et al.*, "Confocal Laser Endomicroscopy as a Guidance Tool for Transbronchial Lung Cryobiopsies in Interstitial Lung Disorder," *Respiration*, 2018.

[74] "Cellvizio System." [Online]. Available: http://www.maunakeatech.com/en/cellvizio/11-cellvizio-system. [Accessed: 05-Feb-2018].